The Evolution of Language

Proceedings of the 12th International Conference on the Evolution of Language (Evolang12)

Editors:

Christine Cuskley Hannah Little Andrea Ravignani Molly Flaherty Luke McCrohon Tessa Verhoef

Evolang 12 Organizing Committee March 13, 2018

Published by:

Evolang 12 Organizing Committee

The Evolution of Language

Proceedings of the 12th International Conference on the Evolution of Language (Evolang12)

Copyright status of individual papers is indicated on their first page. All other sections are licensed under a Creative Commons BY-SA 4.0 license.

Cover photography by Jon Carr, Limor Raviv, Sabrina Engesser, Oren Poliva, Glenn Edward Green, Robert Hawkins, Sonja Vernes and Connie de Vos licensed under a Creative Commons BY-ND 4.0 license.

Evolang XII - the publication of the proceedings has been funded by the Polish Ministry of Science and Higher Education as a part of the 773/P-DUN/2018 agreement



ISBN 978-83-231-3991-1

A digital copy of this volume will remain available for download at: http://www.evolang.org/

In memory of Irit Meir

Preface

The 12th International Conference on the Evolution of Language (EvoLang XII) was held in Toruń, Poland from the 16th - 19th of April 2018, generously organised by the Centre for Language Evolution Studies at Nicolas Copernicus University. The success of the event is due in large part to the dedication of the local organising committee, led by Sławomir Wacewicz and Przemysław Żywiczyński, and with the support of Julia Trzeciakowska and Marek Placiński. Without their contribution, the conference would not have been possible.

The conference received more submissions than ever in this iteration, making the hard work of our reviewers more important than ever. This volume contains well over a hundred high quality contributions from disciplines spanning linguistics, psychology, biology, archaeology, physiology, and genetics (among others). EvoLang XII marks a continued expansion of the EvoLang community, showcasing the ever increasing quality and rigour of empirical and theoretical research on the evolution of language.

This volume would not have been possible without the hard work of the panel of reviewers (p.vii). Following the success of double blind review at EvoLang XI (New Orleans), we implemented the same double-blind procedure this year. The invaluable time and expertise of the review panel, which represents the core of the EvoLang community, is essential to the continued scientific rigour found at EvoLang as we slowly but surely tackle the "hardest problem in science".

As the EvoLang community progresses through its third decade, we now benefit from the expertise of scholars 'raised' on EvoLang from a young (academic) age. Members of the scientific committee have all benefitted in one way or another from mentoring by members of EvoLang's central organising committee (p.vi), who have worked hard to set up several hubs of language evolution research around the world, and who hold the conference together as local and scientific committees come and go.

The EvoLang XII scientific editing committee in particular is a testament to the vital involvement of early career researchers in the evolution of language community, and was a feat of intense international collaboration. The committee this year spanned 3 continents and (at one time or another) 5 countries. Like the local organising committee, the scientific committee voluntarily contribute their time. Thanks are due to my fellow editors for their considerable effort in ensuring the conference not only maintains, but increases its quality with every year.

- Christine Cuskley

Central Organizing Committee

- Andreas Baumann
- Rudolf Botha
- Christine Cuskley
- Erica Cartmill
- Jean-Louis Dessalles
- Ramon Ferrer i Cancho
- Tecumseh Fitch
- Jim Hurford
- Simon Kirby
- Chris Knight

- Heidi Lyn
- Luke McCrohon
- Kazuo Okanoya
- Thom Scott-Phillips
- Andrea Ravignani
- Nikolaus Ritt
- Kenny Smith
- Maggie Tallerman
- Natalie Uomini

Panel of Reviewers

Ann Senghas Anneliese Kuhle Eva Zehentner Joana Rosselló Luis Miguel Rojas-Berscia Matz Larsson Monica A. Irimia Quentin Feltgen Sabrina Engesser Christian Abry Natasha Abner Ines Adornetti Michael Arbib John Archibald David Ardell Rie Asano Mark Atkinson Rachael Bailes Jaume Baixeries Lluís Barceló-Cobliin Mark Bartlett Deniz Baskent Andreas Baumann Ted Bayne André Bazzoni Christina Behme Teresa Bejarano Christian Bentz Aleksandrs Berdicevskis Katrien Beuls Damian Blasi Richard Blythe Bart de Boer Monika Boruta James Brand Svetlana Burlak Ivan Capeller Christophe Coupe Carlos Moreno Juan Cabrera Leandro Caporale Fausto Carcassi Jon Carr Carel ten Cate Alexandre Celma-Miralles Yuan-Lu Chen Alessandra Chiera Katie Collier Dan John Cook Frederick L. Coolidge Marie Coppola Guido Cordoni Davide Crepaldi Jennifer Culbertson Chris Cummins Christine Cuskley Rick Dale Dan Dediu Didier Demolin Daniel Dor Natalia Drozhashchikh Edwige Dugas Michael Dunn Pierce Edmiston T. Mark Ellison Rabia Ergin Livio Favaro Andrew Feeney Olga Feher

Vanessa Ferdinand Ramon Ferrer-i-Cancho Francesco Ferretti Piera Filippi Molly Flaherty Stella Frank Haruka Fujita Koji Fujita Marco Gamba Juan Olvido Perea García Chris Golston Nathalie Gontier Tao Gong Tomás Goucha Jack Grieve Cristina Guardiano Callum Hackett Harald Hammarström Stefan Hartmann Raphaela Heesen Aritz Irurtzun Elizabeth Irvine Rick Janssen Niklas Johansson Sverker Johansson Dariusz Kalociński Jasmeen Kanwal Andres Karjus Fiona Kirton Hiroki Koda Madison Ski Krieger Adriano R. Lameira Marion Laporte Lattenkamp David A. Leavens Ryan Lepic Shiri Lev-Ari Simon D. Levy Molly Lewis David Lightfoot Hannah Little Garv Lupvan Erkki Luuk Heidi Lyn Elainie Madsen Michiru Makuuchi Konstantina Margiotoudi Pedro Tiago Martins Luke McCrohon Adrien Meguerditchian Jérôme Michaud Ashley Micklos Gregory Mills Kivofumi Mochizuki Scott Moisik Kathryn Montemurro Richard Moore Steven Moran Hope Morgan Olivier Morin Yasamin Motamedi Elliot Murphy Thomas Müller Savithry Namboodiripad Dillon Niederhut Alan KS Nielsen Jonas Nölle Tim O'Donnell Cathleen O'Grady

Thomas O'Rourke Mieko Ogura Kazuo Okanova Gerardo Ortega Michał B. Paradowski Joel Parthemore Irene M. Pepperberg Marcus Perlman Michael Pleyer Maciej Pokornowski Ljiljana Progovac Peter Racz Sonia Ragir Andrea Ravignani Limor Raviv Jessica Redmiles Nikolaus Ritt Gareth Roberts Sean Roberts Hanna Ruch Konrad Rudnicki Wendy Sandler Marieke Schouwstra Thom Scott-Phillips Uwe Seifert Yoshimasa Seki Arturs Semenuks Robert Seyfarth Marta Sibierska Catriona Silvey Katie Slocombe Andrew Smith Kenny Smith Göran Sonesson Matthew Spike Michael Spranger Cory Stade James Steele Asha Stewart Justin Sulik Márton Sóskuthy Miki Takahasi Masanori Takezawa Maggie Tallerman Monica Tamariz Bill Thompson Peeter Tinits Fabio Trecca Remi van Trijp Rob Truswell Rory Turnbull Olga Vasileva Freek Van de Velde Tessa Verhoef Sonia Vernes Elisabetta Versace Paul Vogt Connie de Vos Sławek Wacewicz Andrew Wedel James Winters Marieke Woensdregt Martin Zettersten Meng-Han Zhang Qing Zhang Jordan Zlatev Przemyslaw Zywiczynski

Contents

Preface	v
Plenary Talks	1
Iconicity in sign language	2
Children are both more conservative and more likely to gener- alize than adults for the same reason	3
Music before language: Observations from a hunter-gatherers point of view Jerome Lewis	4
Human and pre-human culture and the evolution of language And rew Whiten	5
Refereed Contributions	7
Speech acts addressed at Hadza infants in Tanzania Monika Abels, Paul Vogt	8
Towards a domain-relevant approach to the evolution of lan- guage and music	12
Linguistic and non-linguistic correlates in the evolution of phono- tactic diversity Andreas Baumann, Theresa Matzinger, Nikolaus Ritt	15
Linguistic stability increases with population size, but only in stable learning environments	19

Can Languages without writing systems provide new insights in language evolution? <i>Christina Behme</i>
Language family trees reflect geography and demography be- yond neutral drift Christian Bentz, Dan Dediu, Annemarie Verkerk, Gerhard Jäger
Do non-native speakers create a pressure towards simplification? Corpus evidence
A critical period for the evolution of language-readiness: Clari- fying the Globularization hypothesis <i>Cedric Boeckx</i>
Modelling the Effect of Iconicity on Iterated (Cross-Situational) Learning Bart de Boer
Gestures reflected - Tracing gestural development from a child's EEG signal
Trends in evolution of signals' interpretation as precursors of the origin of human language
Interactive sequences modulate the selection of expressive forms in cross-signing
Conceptual structure is shaped by competing pressures for sim- plicity and informativeness Jon W. Carr, Kenny Smith, Jennifer Culbertson, Simon Kirby
Iconicity and convention in the manual modality: Pantomime in language origins <i>Erica Cartmill</i>
Beat perception in a non-vocal learner: Rats can identify isochronou beats

How domain-specific is merge? Dan Cook, Mauricio Martins	80
Statistical learning and language (in spite of arbitrariness) Davide Crepaldi, Simona Amenta, Marco Marelli	83
Alien symbols for alien language: Iterated learning in a unique, novel signal space <i>Christine Cuskley</i>	86
Cultural transmission of melodic and rhythmic universals: 4 experiments and a model Tania Delgado, Andrea Ravignani, Tessa Verhoef, Bill Thompson, Thomas Grossi, Simon Kirby	89
Phylogeny in phonology: Tai sound systems encode their evo- lutionary history <i>Rikker Dockum</i>	92
Metaphorical Extension and the Evolution of Configurationality T. Mark Ellison, Uta Reinöhl	95
From Cognition to Language Change: Close Simulation of a PNG Village <i>T. Mark Ellison, Eri Kashima</i>	98
Meaning-refining acoustic variation within the internal struc- ture of pied babbler recruitment cries Sabrina Engesser, Amanda Ridley, Marta Manser, Simon Townsend	101
Structural cues for symmetry, asymmetry and non-symmetry in central Taurus sign language Rabia Ergin, Ann Senghas, Ray Jackendoff, Lila Gleitman	104
Dual-Processing and the Representational Hypothesis: Account- ing for the emergence of language Andrew Feeney	107
Semantic conditioning in interaction and transmission Olga Feher, Elizabeth Wonnacott, Hanna Jarvinen, Kenny Smith	111
The coevolution of data and hypotheses in bayesian cultural evolution	114

The evolution of optimized language in the light of standard information theory Ramon Ferrer-I-Cancho, Chris Bentz	122
Do we see word order patterns from silent gesture studies in a new natural language?	125
A rational model of linguistic accommodation and its potential role in language simplification Stella Frank, Kenny Smith	128
Non-communicative functions can be equally important for stud- ies of language evolution	131
Primate songs and their relevance in the study of language evo- lution	134
How the evolution of teaching may help understanding the evo- lution of language <i>Peter Gärdenfors</i>	137
Pattern similarity in biological, linguistic, and sociocultural evo- lution	139
Human and porcine vocalizations after creation of a human larynx Kari Green, Robert Morrison, Glenn Green	150
Natural selection in the modern english lexicon Jack Grieve	153
Learning to communicate about conceptual hierarchies Robert Hawkins, Kenny Smith, Michael Franke, Noah Goodman	158
Linguistic laws in chimpanzee gestural communication Raphaela Heesen, Catherine Hobaiter, Ramon Ferrer-I- Cancho, Stuart Semple	161
Communication of cognition? A minimal computational model of evolutionary interactions between individual and social learn-	
ing Masahiko Higashi, Reiji Suzuki, Takaya Arita	165

When is a code not a code? Elizabeth Irvine	168
Agent model reveals the influence of vocal tract anatomy on speech during ontogeny and glossogeny <i>Rick Janssen, Dan Dediu, Scott Moisik</i>	171
Triangulating sound symbolism where to find it and how to create it	175
Clues to language evolution from a massive dataset with typol- ogy, phonology, and vocabulary from many languages Sverker Johansson	178
Environmental constraints in the evolution of scalar concepts: Road to most Dariusz Kalociński	181
Identifying linguistic selection and innovation while controlling for cultural drift	189
Temporal adjustment of short calls during vocal exchange in Japanese macaques Noriko Katsu, Kazunori Yamada, Masayuki Nakamichi, Kazuo Okanoya	193
Learning implicational models of universal grammar parameters Dimitar Kazakov, Guido Cordoni, Andrea Ceolin, Mon- ica Irimia, Shin-Sook Kim, Dimitris Michelioudakis, Nina Radkevich, Cristina Guardiano, Giuseppe Longobardi	196
The role of teaching in iterated language transmission Vera Kempe, Kamil Cichon, Monica Tamariz, Nicolas Gauvrit	206
Incremental word processing helps shape the lexicon	209
The origins of word order universals: Evidence from corpus statistics and silent gesture Simon Kirby, Jennifer Culbertson, Marieke Schouwstra	212
Constituent Order in Improvised Gesture Reflects Speaker Per- spective Fiona Kirton, Marieke Schouwstra, Jennifer Culbertson, Kenny Smith, Simon Kirby	215

Pressure for trust-based efficiency shaped the evolution of lan- guage Chris Knight	218
Do children use "fine" ostensive communication? Harumi Kobayashi, Tetsuya Yasuda	230
Recontextualization: The Dynamics of Language Behavior and Change Anneliese Kuhle, Ferdinand von Mengden	232
Mammalian models for the study of vocal learning: A new paradigm in bats Ella Lattenkamp, Sonja Vernes, Lutz Wiegrebe	235
Sequencing bat genomes to shed light on language evolution Ksenia Lavrichenko, Sonja Vernes	238
The Method of Validation by Zenith David Leavens	241
What 50 Million Drawings Can Tell Us About Shared Meaning Molly Lewis, Gary Lupyan	243
What do iconicity judgements really mean? Hannah Little, Justin Sulik	246
Physical features of social networks and language typology Matthew Lou-Magnuson, Luca Onnis	249
Core Knowledge or language-augmented cognition? The case of geometric reasoning Gary Lupyan, Ashley Wendorf, Luis Miguel Berscia, Jing Paul	252
Gesture, sign and beyond. Negation across three generations of signers	255
Symbol and its evolution Erkki Luuk	258
Greeting and meeting: Vocal interactions and use of signature whistles in bottlenose dolphins during an introduction <i>Heidi Lyn, Megan Broadway</i>	270

Neuroimaging studies on drawing suggest a link between sym- bolic behavior to language: A potential hint to the Nean- derthal language controversy	273
What can sound symbolism tell us about the evolution of lan- guage? A comparative study in humans and great apes Konstantina Margiotoudi, Manuel Bohn, Matthias Allritz, Friedemann Pulvermüller	278
The cultural evolution of communicative conventions: Interac- tions between population connectivity dynamics and cognitive biases	281
Sound Production Learning and the Contiguum Hypothesis Pedro Tiago Martins, Cedric Boeckx	284
The human Arcuate Fasciculus provides specific advantages to process complex sequential stimuli, not hierarchies in general <i>Mauricio Martins, Arno Villringer</i>	287
A potential link between infant-directed speech and evolution of vocal communication <i>Reiko Mazuka, Tadahisa Kondo, Akiko Hayashi</i>	290
Identifying the Ape Beat in the Wild: Rhythmic Individual Sig- natures from the Sounds of Manual Fruit Cracking in Fongoli Chimpanzees	293
Topic-openendedness: Why recursion is overrated Irit Meir	295
Dynamical preferences and self-actuated language changes \dots Jérôme Michaud	306
The prevalence of repair in studies of language evolution Ashley Micklos, Vinicius Macuch Silva, Nicolas Fay	316
The emergence of adjacency pairs: No evidence is better than negative evidence	319
Differential rates of change in consonant and vowel systems Steven Moran, Annemarie Verkerk	322

The cultural evolution of spatial modulations in artificial sign languages	326
Evolutionary Stability of Linguistic Politeness and the Polite- ness Equilibrium Principle	329
The influence of shared visual context on performance with novel codes in a referential communication game Thomas Müller, Tiffany Morisseau, James Winters, Olivier Morin	332
A domesticated code: On the emergence of the oscillatory basis of phrase structure <i>Elliot Murphy</i>	335
Semantic bleaching not observed in synchronic test Dillon Niederhut	339
Comparative morphology of the laryngeal muscles in hylobatids using a high-resolution MRI Takeshi Nishimura, Yoshitaka Nomura, Hirohiko Imai, Tet- suya Matsuda	346
Environmental and social factors motivate the emergence of sys- tematic categories and signs Jonas Nölle, Marlene Staib, Riccardo Fusaroli, Kristian Tylén	349
The evolutionary causes of irregularity: Evidence for an irreg- ularization bias in morphological learning <i>Timothy O'Donnell, Kenny Smith</i>	352
Evolution of Homophones and Syntactic Categories Noun and Verb Mieko Ogura, William Wang	355
Language as a parasite to emotional signals	364
Socio-environmental Constraints in Category Learning Katarzyna Oktaba, Dariusz Kalociński	367
Simplicity biases operate at item and system level. An iterated learning study	370

A conceptual escape from the perils of universal grammar Joel Parthemore	372
Evolution of phoneme usage driven by phonetic robustness for error minimization Ines Peña, Marco Archetti	376
The evolutionary dynamics of language as a function of demog- raphy Dirk Pijpops, Katrien Beuls, Freek Van de Velde	379
Language evolution and the emergence of constructions in in- teraction	382
From where to what: A neuroanatomically based evolutionary model of the emergence of speech in humans <i>Oren Poliva</i>	386
Pinnipeds have something to say about speech and rhythm Andrea Ravignani, Maxime Garcia, Stephanie Gross, Koen de Reus, Nienke Hoeksema, Ana Rubio García, Bart de Boer	399
The role of community size in the emergence of linguistic structure $Limor\ Raviv,\ Antje\ Meyer,\ Shiri\ Lev-Ari$	402
Conceptual categories scaffold verbal semantic structure: A cross-cultural study of child homesign Lilia Rissman, Laura Horton, Susan Goldin-Meadow	405
Word form shapes are culturally selected for indicating their morphological structure	408
CHIELD: Causal hypotheses in evolutionary linguistics database $Sean\ Roberts$	412
Innovation, selection and the emergence of transparent signals in interaction	419
The emergence of focus in an experimental communication game $Gareth\ Roberts,\ Jon\ Stevens$	422
Compressible kinship terminologies are more learnable than less compressible alternatives	425

Homo imitans is homo loquens. On vocal imitation Joana Rossello	428
The effects of gossip on interpersonal trust depend on proso- ciality of the gossipers	431
Distribution-based Prediction of the Degree of Grammaticaliza- tion for German Prepositions Dominik Schlechtweg, Sabine Schulte Im Walde	434
Cockatiels: A novel animal model for studying the evolution of music and language Yoshimasa Seki	437
What Makes a Grammar Difficult? Experimental Evidence Arturs Semenuks, Aleksandrs Berdicevskis	439
Evidence for High Frequency of Linguistic Rara In European Languages Arturs Semenuks	442
Different patterns of iconic influence in the creation vs. trans- mission of the Nicaraguan sign language lexicon Ann Senghas, Jennie Pyers, Caroline Zola, Charlotte Quin- coses	445
Abstraction and universals in the orizing about language evolution $Richard\ Shillcock$	449
Structural alignment in cross-modal priming: Linguistic repre- sentation is shared between gesture and speech Joleana Shurley, Marieke Schouwstra, Martin Pickering	458
The Informative Value Of Communicator Disagreement in Great Apes For The Study Of The Evolution Of Human Intentional Communication Christine Sievers, Thibaud Gruber, Klaus Zuberbühler	461
'Give it to me'/'Give it to her': The development of a relational term in homesign and English <i>Catriona Silvey, Susan Goldin-Meadow</i>	464
Referential gestures are not ubiquitous in wild chimpanzee com- munities	467

Does learning favour communicative efficiency? Kenny Smith, Jennifer Culbertson	469
Understanding vs. Describing Others' Minds: Mental State Language in a Small-Scale Society Andrew Marcus Smith, Erica Cartmill, H. Clark Barrett	472
How sociolinguistic theory can inform cultural evolution, and how to test it	475
The greatest story ever told: Semiosis emerging from mimesis and/or narrativity <i>Göran Sonesson</i>	478
Language complexity as an interaction between social structure, innovation, and simplicity	489
I see what you did there: The role of iconicity in the acquisition of signs	492
Success in signaling: the effect of feedback to signaler and receiver Justin Sulik, Gary Lupyan	495
Modeling creativity and communication Justin Sulik	498
Development of vocal temporal parameters in distantly related vocal learners, birds and humans <i>Miki Takahasi, Kazuo Okanoya, Reiko Mazuka</i>	501
Do songbirds perceive syllables as isolated tokens? Mizuhara Tomoko, Kazuo Okanoya	504
Difference of two recursions, "hierarchical embedding" and "self- reference": From view of adaptive functions and implementa- tions	507
Communicative asymmetry, handedness and language experience $Olga\ Vasileva$	510
Cognitive biases and cultural evolution in the emergence of space-time mappings in language Tessa Verhoef, Esther Walker, Tyler Marghetis	513

Vocal learning in bats: From genes to behaviour Sonja Vernes, Ella Lattenkamp, Ksenia Lavrichenko, Ja- nine Mengede	516
Honeybees learn patterns with flashing lights Elisabetta Versace, Sara Peruffo, Bettina Maccagnani, Giorgio Vallortigara	519
Rapid turn-taking as a constant feature of signed conversations Connie de Vos	522
Phonological grammars evolve to preserve information at word beginnings Andy Wedel, Adam Ussishkin, Adam King, Jaycie Martin, Jonathon Geary	525
From context to code: Shared information constrains the evo- lution of optimal codes James Winters, Olivier Morin	528
A model of cultural co-evolution of language and perspective- taking Marieke Woensdregt, Kenny Smith, Chris Cummins, Si- mon Kirby	531
Infants' sensitivity to emotional animal vocalization and the evolution of vocal communication Naoto Yamane, Mihoko Hasegawa, Ai Kanato, Naoko Ki- jima, Kazuo Okanoya, Reiko Mazuka	534
The English dative alternation as an adaptation to changes in the constructional network	537
Using language to discover categories: More nameable features improve category learning	539
Pantolang: A synthetic cognitive-semiotic approach to language origins Jordan Zlatev, Simon Devylder, Slawomir Wacewicz, Prze- myslaw Zywiczynski, Francesco Ferretti, Ines Adornetti, Alessandra Chiera	548
Vocal accommodation in common marmosets: Does similarity buffer tension during pair bond development? Yvonne Zürcher, Erik P. Willems, Judith M. Burkart	551

Author Index

555

Plenary Talks

ICONICITY IN SIGN LANGUAGE

KAREN EMMOREY^{*1}

*Corresponding Author: kemmorey@mail.sdsu.edu ¹San Diego State University, California, USA

Iconicity may have played a role in language evolution under the assumption that the earliest vocal or manual communications involved motivated forms. The study of sign languages may provide some insight into this issue because the visual-manual modality allows for iconic expression of a wide range of conceptual elements, e.g., object and human actions, movements, locations, and shapes. Currently, however, we know very little about how visual-manual iconicity is perceived by signers vs. non-signers or whether iconic signs are processed differently in the brain. My colleagues and I have been exploring the nature of the distribution of iconic forms in the American Sign Language (ASL) lexicon, how the perception of iconicity is impacted by linguistic knowledge, and how the perception of iconicity changes when the sign meaning is given versus when it must be guessed. We have also been investigating whether there are general principles of alignment between visual-manual articulations and conceptual elements (e.g., handshapes map to objects). Finally, we have been using Event-Related Potentials (ERPs) to investigate the possible role of iconicity in modulating the temporal neural dynamics of single sign processing. Thus far, the ERP results indicate that iconic signs do not have a distinct neural signature, but frequency and concreteness effects are parallel for word and sign recognition. Overall, this work indicates a) an important distinction between iconicity and transparency, b) linguistic knowledge reduces and changes sensitivity to iconicity, and c) there appears to be no distinct neural response to iconic signs.

CHILDREN ARE BOTH MORE CONSERVATIVE AND MORE LIKELY TO GENERALIZE THAN ADULTS FOR THE SAME REASON

ADELE GOLDBERG^{*1,}

*Corresponding Author: adele@princeton.edu ¹Psychology Department, Princeton University, New Jersey, USA

Children have been argued to be both more conservative than adults, initially appearing to be less productive with their linguistic constructions than adults; they have also been argued to be over-eager generalizers when compared with adults, insofar as they are more likely to regularize complex linguistic input. This raises an apparent paradox: how can children be more likely than adults to both under-generalize (be conservative) *and* over-generalize (regularize)? The answer lies in the fact that children are less adept at aligning exemplars within the high-dimensional conceptual space in which our long-term knowledge of language—our *constructicon*—is represented. Children are therefore more prone than adults to be conservative, when they fail to recognize relevant similarities among exemplars, and simultaneously more likely to oversimplify or generalize, when they fail to appreciate relevant distinctions.

MUSIC BEFORE LANGUAGE: OBSERVATIONS FROM A HUNTER-GATHERER'S POINT OF VIEW

JEROME LEWIS^{*1}

*Corresponding Author: jerome.lewis@ucl.ac.uk ¹Department of Anthropology, University College London, London, UK

If language evolved when humanity lived in Africa as hunter-gatherers, then what might African hunter-gatherers have to tell us about the evolutionary relationship between music and language? The BaYaka's view of the human communicative spectrum extends from using plants to signal, signing and animal mimicry, through a range of human languages and linguistic devices, to the employment of music, dance and other performative strategies. Unfettered by standardisation or identity concerns, BaYaka are communicative predators, prioritising efficacy in their techniques for communicating with a diversity of human and non-human sentient and responsive elements of their environment. Different but complementary, these modes of communication are designed for a range of audiences. At one end, signing and speech are for an individual to communicate to others, while singing in musicking groups is intended to facilitate group level communication at the other.

Taking a reverse anthropology approach, Bayaka women's reasons for singing suggest a behavioural ecology explanation for the prior emergence of music. The ethnography suggests that musical ritual sufficiently respects costly signalling constraints to provide the key scaffolds required for language: capacities for vocal dexterity and vocal learning while participation stimulates we-intentionality, so enabling normativity to develop. These are key building blocks for the emergence of language.

HUMAN AND PRE-HUMAN CULTURE AND THE EVOLUTION OF LANGUAGE

ANDREW WHITEN*1

*Corresponding Author: aw2@st-andrews.ac.uk ¹University of St Andrews, Fife, Scotland, UK

Human language is intimately linked with human culture. A child's language is culturally learned from others; and conversely, language can be a major vehicle of cultural transmission. To understand the evolution of language, we therefore need to understand the evolution of the capacity for culture. In this we are lucky, for by contrast with the apparent gulf between human language and all other animal communication, I shall describe how culture in some very recognisable forms has been discovered to be surprisingly widespread among animals and particularly rich in our closest primate relatives, from which we can make inferences about the ancestry of our cultural cognition. Tracing the evolution of culture is thus arguably a much more tractable empirical quest than that of language per se, and in this talk I will review research by my own and other research groups that pursues this through a multitude of different and convergent methods, from ethological observation in the wild to a diversity of cultural diffusion experiments, completed with both human and non-human primates in our own research. Mindful of the debate that pits a gestural against a vocal origin hypothesis for the evolution of language (the meeting of criteria for intentional communication argued to favour the gestural candidate), I note there is much evidence for vocal learning amongst varied animal taxa (although relatively little amongst primates) compared with little or none for the cultural transmission of gesture. This may suggest that vocal learning from others is compared to copying gestural communication. Yet, 'easy' perhaps paradoxically, the evidence for cultural transmission in primates is massively concerned with manual actions, such as those directed at functional, material outcomes, like tool use. I conclude that while they lack much sign of linguistictype communication, contemporary apes, and hence likely our early hominin

ancestors, would be well prepared in capacities for cultural transmission once such forms of communication emerged. I describe evidence for several features of ape social learning that suggest pre-adaptations for transmission of aspects of linguistic communication, including copying 'syntactical' sequential and hierarchical structure of actions, 'rational' imitation, cumulative learning and reflexive recognition of imitation.

Refereed Contributions

SPEECH ACTS ADDRESSED AT HADZA INFANTS IN TANZANIA

MONIKA ABELS^{*1} and PAUL A. VOGT¹

^{*}Corresponding Author: monika.abels@gmail.com ¹Tilburg School of Humanities and Digital Sciences, Tilburg University, Netherlands

1. Introduction

Language socialization is known to differ between different human societies (e. g. Ochs & Schieffelin, 1984). "Western" societies foster cognitive skills (e.g., language), but in many sub-Sahara African, rural societies (communal) action autonomy is considered more important (Keller, 2011). Different parental beliefs appear related to the speech acts addressed to children: Western, middle class caregivers tend to use more declaratives and questions when addressing their infants, whereas in many non-Western, rural communities caregivers tend to use relatively more imperatives (Rabain Jamin & Sabeau-Jouannet, 1997; Vogt, Mastin, & Schots, 2015).

Early language socialization has rarely been studied among hunter-gatherer communities. From an evolutionary point of view, studying the language socialization of hunter-gatherers could provide crucial insights into interactional settings that may have been similar for our ancestors (Marlowe, 2010). The egalitarian culture of hunter-gatherer societies (Marlowe, 2010) may influence the type of speech acts caregivers address to infants in unexpected ways. For example, it has been reported that hunter-gatherer infants receive few direct instructions, which could affect the amount of imperatives addressed to infants (Hewlett & Roulette, 2016).

The Hadza are traditionally hunter-gatherers in northern Tanzania. Recent changes in livelihood can be observed in present day Hadza camps. For example more tourists visit some of the camps bringing money to pay for tour guides and souvenirs. These changes impact the Hadzas' lifestyle, and consequently may also influence the language socialization of their infants.

The following research questions are addressed in this study: (1) Are traditional child rearing ideals (i.e., little explicit instruction) reflected in caregivers' speech acts? (2) Do speech acts differ with camp livelihood?

2. Methods

Twenty-five infants of 6 - 27 months were visited in their camps with different livelihoods ((almost) daily contact with tourists (n=7), occasional/indirect tourism or farming (n=13), isolated (n=5)). After familiarization, infants were video-recorded on average 113 minutes (range 33-176) during their normal daily activities such as playing and eating. All videos were transcribed and translated to English by a native speaker of *Hadzabe*. Speech acts were coded as described by (Rabain-Jamin, 2001) as "Assertives", "Requests for Information" and "Requests for Action" and additionally "Vocatives" (Van de Weijer, 1999).

3. Results and Discussion

The results (cf. Appendix) suggest that Hadza infants experience speech acts similar to those of other infants in sub-Saharan Africa (Rabain-Jamin, 2001; Vogt et al., 2015), with very few assertives and requests for information, but frequent requests for actions. However, in isolated camps caregivers use significantly fewer requests for actions and relatively more vocatives than the less traditional camps. These findings suggest that traditional Hadza speech acts may have adhered to hunter-gatherer child rearing practices consisting of fewer imperatives than may be found in rural communities or less traditional huntergatherer communities, but considerably more than in Western communities. Furthermore, the low amount of assertives or requests for information indicates that the fostering of cognitive skills is not considered crucial, although this does not suggest that hunter-gatherers have no teaching strategies as is sometimes suggested (cf. Hewlett & Roulette, 2016). It could also indicate that early human speech was not assertoric, as suggested by Tomasello (2010) but may have fulfilled function such as coordination and participation (Rappaport, 1999). The relatively high frequency of vocatives suggests an early fostering of relatedness (Biber et al., 1999), which may indicate a crucial role of communicating relations between people (Fitch, 2004) or assuring infants of caregivers' presence (Falk, 2004) during the early language evolution.

Acknowledgements

We thank all participating infants and caregivers. We are indebted to Ruth Matiyas, Mariamu Anyawire, Gudo Mahiya and Bernadetha Sanka for their help in data collection, transcription and translation. We are grateful to Andrew Kilale and John Ogondiek from the Tanzanian National Institute of Medical Research for their support. We thank Camilla Power from the University of East London for her help in planning and organizing this research and her input on Hadza culture. This project has received funding from the European Union through a Marie Skłodowska-Curie Actions Individual Fellowship (H2020-MSCA-IF-2014) for the project "Hadza Communication".

References

- Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E., & Quirk, R. (1999). Longman grammar of spoken and written English (Vol. 2). Cambridge, MA: MIT Press.
- Falk, D. (2004). Prelinguistic evolution in early hominins: whence motherese? Behav Brain Sci. 27 (4): 491–503.
- Fitch, W. T. (2004). Kin selection and "mother tongues": a neglected component in language evolution. In U. Griebel & D. K. Oller. Evolution of communication systems : a comparative approach. Cambridge, Mass.: MIT Press. pp. 275–296.
- Hewlett, B. S., & Roulette, C. J. (2016). Teaching in hunter–gatherer infancy. *Royal Society Open Science*, 3(1), 150403. http://doi.org/10.1098/rsos.150403
- Keller, H. (2011). Autonomy and Relatedness Revisited: Cultural Manifestations of Universal Human Needs. *Child Development Perspectives*, 6(1), 12–18. http://doi.org/10.1111/j.1750-8606.2011.00208.x
- Marlowe, F. (2010). The Hadza: Hunter-Gatherers of Tanzania (Origins of Human Behavior and Culture) (1st ed.). University of California Press.
- Ochs, E., & Schieffelin, B. (1984). Language acquisition and socialization: Three developmental stories and their implications. In R. Shweder & R. Levine (Eds.), *Culture theory: essays on mind, self and emotion* (pp. 276– 320). New York: Cambridge University Press.
- Rabain Jamin, J., & Sabeau-Jouannet, E. (1997). Maternal speech to 4-monthold infants in two cultures: Wolof and French. *International Journal of Behavioral Development*, 20(3), 425–451. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/016502597385216
- Rabain-Jamin, J. (2001). Language use in mother-child and young sibling interactions in Senegal. *First Language*, 21(63), 357–385. http://doi.org/10.1177/014272370102106307
- Rappaport, R. (1999). *Ritual and religion in the making of humanity*. Cambridge, U.K. New York: Cambridge University Press.
- Tomasello, M. (2010) Origins of human communication. MIT press.
- Van de Weijer, J. C. (1999, January 11). Language Input for Word Discovery.

(A. Cutler, Ed.). Katholieke Universiteit Nijmegen.

Vogt, P., Mastin, J. D., & Schots, D. M. A. (2015). Communicative intentions of child-directed speech in three different learning environments: Observations from the Netherlands, and rural and urban Mozambique. *First Language*, 35(4-5), 341–358. http://doi.org/10.1146/annurev.an.15.100186.001115

Appendix - statistics

Speech act	F (2, 22)	р	$ \begin{array}{c} partial \\ \eta^2 \end{array} \end{array}$	Tourist		Transitional		Isolated		LSD Post-
				mean	SD	mean	SD	Mean	SD	hoc
Proportions										
Request for information	1.99	.161	.15	7.14	5.39	5.38	3.85	2.20	3.38	-
Request for action	8.07	.002	.42	74.54	9.26	60.79	11.19	51.70	7.28	To > Tr = Is
Assertive	4.00	.033	.27	11.59	8.25	21.05	7.50	15.41	4.89	Tr > To
Vocative	12.75	<.001	.54	6.73	5.88	12.78	8.40	30.69	10.88	Is > Tr = To
Total number of speech acts/minute				0.80	0.52	0.75	0.47	1.28	0.49	

Differences in speech acts by camp type

TOWARDS A DOMAIN-*RELEVANT* APPROACH TO THE EVOLUTION OF LANGUAGE AND MUSIC

RIE ASANO^{*1}

*Corresponding Author: rie.asano@uni-koeln.de ¹Department of Systematic Musicology, University of Cologne, Cologne, Germany

The central idea of current comparative research on the evolution of language and music is that they consist of multiple components with different evolutionary origins (Fitch, 2006). From a comparative language-music perspective, some components might be shared and based on the same evolutionary genesis, while others might be different and emerged independently in the course of evolution. From a comparative between-species perspective, some might be shared with other animals, while others might be unique to humans. However, this shared-distinct dichotomy dominating the recent comparative approach usually depends on tailor-made categories fitting to just one domain or species and thus limits the range of investigation by its all-or-nothing contrastive view (De Waal & Ferrari, 2010; Theofanopoulou & Boeckx, 2015).

To go beyond such a shared/distinct dichotomy, the current paper puts a domain-*relevant* approach forward. Its main idea is that through neural competition, brain networks become relatively domain-specific overtime (Karmiloff-Smith, 2013). Thus, specialization of function can be regarded as fine tuning of coarsely coded systems with domain-relevant biases. In evolutionary research, those domain-*relevant* systems should be the target of comparative endeavor investigating the evolution of language and music. Based on evidence from cognitive and evolutionary neuroscience (Kotz et al., 2009; Lieberman, 2016; Merchant et al., 2015; Ullman, 2006) as well as modeling studies (e.g., Dominey et al., 2009), I suggest that the cortico-basal ganglia-thalamocortical (CBGT) circuits form a promising candidate for such systems.

The CBGT circuits are involved in and are necessary for performing a procedure, i.e. organizing sequences of actions towards a goal. For example, syntactic sequence processing in language and beat-based rhythmic sequence

processing in music build on those circuits. They are also required for procedural learning such as habit and rule learning. Impairments of the CBGT circuits result in underspecified rule representations in language and music. Therefore, from a neurocognitive perspective the cognitive systems language and music can be regarded as a different use of the same domain-*relevant* systems.

Moreover, investigations of the CBGT circuits provide direct betweenspecies comparative options: those circuits are required for non-human primates' action cognition (Mendoza & Merchant, 2014) as well as song learning in songbirds (Jarvis, 2004). Thus, the current paper provides strong support for hypotheses that regard both these current neurocognitive systems as products of evolutionary changes of an ancestral action cognition systems (Boeckx & Fujita, 2014; Fujita, 2016; Marcus et al., 2006).

Acknowledgements

Special thanks to Prof. Uwe Seifert for his helpful comments to the earlier version of this abstract.

References

- Boeckx, C., & Fujita, K. (2014). Syntax, action, comparative cognitive science, and Darwinian thinking. *Frontiers in Psychology*, *5*, 627.
- De Waal, F. B. M., & Ferrari, P. F. (2010). Towards a bottom-up perspective on animal and human cognition. *Trends in Cognitive Sciences*, 14(5), 201–7.
- Dominey, P. F., Inui, T., & Hoen, M. (2009). Neural network processing of natural language: II. Towards a unified model of corticostriatal function in learning sentence comprehension and non-linguistic sequencing. *Brain and Language*, 109(2–3), 80–92.
- Fitch, W. T. (2006). The biology and evolution of music: a comparative perspective. *Cognition*, *100*(1), 173–215.
- Fujita, K. (2016). On the parallel evolution of syntax and lexicon: A Merge-only view. *Journal of Neurolinguistics*, 1–15.
- Jarvis, E. D. (2004). Learned Birdsong and the Neurobiology of Human Language. *Annals of the New York Academy of Sciences*, 1016(1), 749–777.
- Karmiloff-Smith, A. (2013). Challenging the use of adult neuropsychological models for explaining neurodevelopmental disorders: developed versus developing brains. *Quarterly Journal of Experimental Psychology*, 66(1), 1– 14.
- Kotz, S. A., Schwartze, M., & Schmidt-Kassow, M. (2009). Non-motor basal ganglia functions: A review and proposal for a model of sensory predictability in auditory language perception. *Cortex*, 45(8), 982–990.

- Lieberman, P. (2016). The evolution of language and thought. *Journal of Anthropological Sciences*, 94, 1–20.
- Marcus, G. F., Rabaglia, C. D., & Rabagliati, H. (2006). Modularity and descentwith-modification. In C. Boeckx & K. K. Grohmann (Eds.), *The Cambridge Handbook of Biolinguistics* (pp. 326–340). Cambridge: Cambridge University Press.
- Mendoza, G., & Merchant, H. (2014). Motor system evolution and the emergence of high cognitive functions. *Progress in Neurobiology*, 122, 73–93.
- Merchant, H., Grahn, J. A., Trainor, L., Rohrmeier, M., & Fitch, W. T. (2015). Finding the beat: a neural perspective across humans and non-human primates. *Phil. Trans. R. Soc. B*, 370, 20140093.
- Theofanopoulou, C., & Boeckx, C. (2015). Cognitive phylogenies, the Darwinian logic of descent, and the inadequacy of cladistic thinking. *Frontiers in Cell* and Developmental Biology, 3(October), 1–6.
- Ullman, M. T. (2006). Is Broca's Area Part of a Basal Ganglia Thalamocortical Circuit? *Cortex*, 42(4), 480–485.

LINGUISTIC AND NON-LINGUISTIC CORRELATES IN THE EVOLUTION OF PHONOTACTIC DIVERSITY

ANDREAS BAUMANN*1, THERESA MATZINGER1, and NIKOLAUS RITT1

*Corresponding Author: andreas.baumann@univie.ac.at ¹Department of English and American Studies, University of Vienna, Vienna, Austria

Linguistic dynamics have been hypothesized to be driven by ecological factors such as population size or social structure (see Nettle, 2012 for an excellent overview). Particularly, there is an ongoing debate as to whether population size can be seen as an explanatory factor in the evolution of phonemic richness (Atkinson, 2011; Bybee, 2011; Hay & Bauer, 2007; Wichmann, Rama, & Holman, 2011; see also Moran, McCloy, & Wright, 2012 for critical discussion). In this regard, the evolution of larger sublexical constituents, i.e. sequences of sounds below the word level, has gained much less attention (but see Maddieson, 2013 or Rama, 2013). Moreover, studies on the connection between ecological factors and linguistic properties were primarily comparative in nature, although the parallel evolution of social structure and language in individual linguistic strands may also provide useful insights into the mechanics that drive language evolution (see Bybee, 2011; Pagel, Atkinson, & Meade, 2007; Trudgill, 2004).

In this paper, we conceptualize phonotactic items (sequences of sounds) as culturally transmitted pieces of linguistic knowledge, i.e. competence constituents in their own right, which spread through populations just like single sounds, words or constructions (Croft, 2000; Ritt, 2004). Phonotactic items should therefore be subject to similar evolutionary pressures and mechanisms. We investigate the diachronic development of diversity of the phonotactic inventory in the history of English from Middle English to Present Day English (using historical data from PPCME2, PPCEME, PPCMBE and COHA, and phonological transcriptions from ECCE and CMU). We focus on word final phonotactics because changes are most likely to occur at this prosodically weak position, and for methodological reasons (fully phonologically analyzed historical texts are not available for early periods). We find that the diversity of word-final coda phonotactics has been increasing through the past 800 years, and that the
evolution of phonotactic diversity is strongly related to that of network characteristics that can be derived from population size.

Our approach goes like this: for each period of 50 years from 1150 to 2000, we computed *true diversity* (cf. Tuomisto, 2010) based on the respective frequency distributions of word-final consonant sequences. The resulting trajectory indicates that English phonotactics became more diverse. We then retrieved trajectories for potentially related features that fall into three categories: (a) linguistic features (size of diphone inventory; consonant-inventory size; syntheticity; analyticity; cf. Szmrecsanyi, 2012), (b) socio-geographic features (population size; populated area; population density of the English speaking community; estimates taken from Wrigley & Schofield, 1981 and more recent census data), and (c) network features directly derived from population size under the assumption of a scale-free small-world network (network diameter; clustering coefficient; Barabási, 2016). In total, this amounts to ten trajectories.

In order to compare the trajectories to each other and to find out which development matches that of phonotactic diversity best, we use autocorrelationdriven time-series clustering (this has – in contrast to e.g. Pearson or Minkowskidistance based procedures – the advantage of also taking the temporal structure into account; see Montero & Vilar, 2014 and references therein). We find that the evolution of phonotactic diversity correlates most strongly with that of the computed clustering coefficient (albeit in a negative way: high clustering corresponding to low diversity) and with the trajectories of population density, populated area and network diameter (all positively correlated). Phonotactic inventory size (i.e. diversity as measured in Rama, 2013) correlates less strongly with factors in that group (which entails as a corollary that dynamics in phonotactic diversity are not just a reflex of increased lexical diversity due to loan import etc.). The remaining trajectories (notably population size together with the other linguistic features) form separate groups.

Our analysis yields a number of insights. First, it suggests that it is probably not population size itself (and associated exposure to drift effects) which directly affects linguistic evolution but rather more immediate (but related) factors that determine the amount and heterogeneity of linguistic interactions. Indeed, increased clustering (i.e. the tendency of forming small groups) has been shown to decrease growth of new variants (Miller, 2009). Likewise, high population density can be argued to promote the spread of linguistic constituents (even if they *a priori* have deficient reproductive properties like sequences of consonants in the prosodically weak coda-position). Second, on a more methodological level, we argue that language-dating methods which are based on phonotactic diversity (Rama, 2013) must take population size and related factors into account in order to prevent the method from just reflecting world-wide population increase in the past centuries. Finally, in agreement with Bybee (2011), we stress that measures of linguistic diversity which also take token frequency into account (such as entropy or true diversity) might be more profitable for the research on language evolution than counts of types (e.g. phoneme or diphone inventory size, lexicon size).

Acknowledgements

We would like to thank Klaus Hofmann for helpful comments on this study. This research was supported by Austrian Science Fund (FWF, grant No. P27592-G18).

References

- Atkinson, Q. D. (2011). Phonemic diversity supports a serial founder effect model of language expansion from Africa. *Science*, 322(6027), 346–349.
- Barabási, A.-L. (2016). *Network science*. Cambridge: Cambridge University Press.
- Bybee, J. (2011). How plausible is the hypothesis that population size and dispersal are related to phoneme inventory size? Introducing and commenting on a debate. *Linguistic Typology*, *15*, 147–153. https://doi.org/10.1515/LITY.2011.009
- Croft, W. (2000). *Explaining language change: An evolutionary approach. Longman linguistics library*. Harlow, England, New York: Longman.
- Hay, J., & Bauer, L. (2007). Phoneme inventory size and population size. Language, 83(2), 388–400.
- Maddieson, I. (2013). Syllable Structure. In M. S. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Max Planck Digital Library. Retrieved from http://wals.info/chapter/12
- Miller, J. (2009). Spread of infectious disease through clustered populations. J R Soc Interface, 6, 1121–1134.
- Montero, P., & Vilar, J. (2014). TSclust: An R Package for Time Series Clustering. *Journal of Statistical Software*, 62(1).
- Moran, S., McCloy, D., & Wright, R. (2012). Revisiting population size vs. phoneme inventory size. *Language*, 88(4), 877–893. https://doi.org/10.1353/lan.2012.0087
- Nettle, D. (2012). Social scale and structural complexity in human languages. *Phil. Trans. R. Soc. B*, 367, 1829–1836.

- Pagel, M., Atkinson, Q. D., & Meade, A. (2007). Frequency of word-use predicts rates of lexical evolution throughout Indo-European history. *Nature*, 449(7163), 717–720.
- Rama, T. (2013). Phonotactic diversity predicts the time depth of the world's language families. *PloS one*. Advance online publication. https://doi.org/10.1371/journal.pone.0063238
- Ritt, N. (2004). Selfish sounds and linguistic evolution: A Darwinian approach to language change. Cambridge: Cambridge University Press.
- Szmrecsanyi, B. (2012). Analyticity and syntheticity in the history of English. In T. Nevalainen & E. C. Traugott (Eds.), *The Oxford handbook of the history* of English (pp. 654–665). Oxford: Oxford University Press.
- Trudgill, P. (2004). Linguistic and social typology: The Austronesian migrations and phoneme inventories. *Linguistic Typology*, 8(3), 305–320.
- Tuomisto, H. (2010). A consistent terminology for quantifying species diversity? Yes, it does exist. *Oecologia*, 164, 53–860. https://doi.org/10.1007/s00442-010-1812-0
- Wichmann, S., Rama, T., & Holman, E. (2011). Phonological diversity, word length, and population sizes across languages: The ASJP evidence. *Linguistic Typology*, 15(2), 177–197.
- Wrigley, E. A., & Schofield, R. (1981). The Population History of England, 1541–1871. A reconstruction. Harvard: Harvard University Press.

LINGUISTIC STABILITY INCREASES WITH POPULATION SIZE, BUT ONLY IN STABLE LEARNING ENVIRONMENTS

ANDREAS BAUMANN^{*1}

*Corresponding Author: andreas.baumann@univie.ac.at ¹Department of English and American Studies, University of Vienna, Vienna, Austria

The effect of population size on linguistic stability and evolution has been investigated in different linguistic domains. The relationship among these factors, however, is not always clear. In this paper, we study a basic population-dynamical model of linguistic spread, derive measures of linguistic stability and fitness, and investigate the effect of population size on these measures. By allowing for stochasticity in the learning process of linguistic constituents, it is shown that a constituent's stability and fitness increases with population size, but that high variability in the learning environment may cause constituent loss, also in large populations. The respective roles of learning and usability are also discussed.

1. Population size and linguistic evolution

Population size has been proposed to affect linguistic structure (Atkinson, Kirby, & Smith, 2015; Hay & Bauer, 2007; Lupyan & Dale, 2010; Nettle, 2012; Wichmann, Rama, & Holman, 2011) as well as rate of linguistic change (Atkinson, 2011; Wichmann & Holman, 2009) and degree of adaptation with respect to cognitive and communicative pressures (Fay & Ellison, 2013). More recently, Bromham et al. (2015) have shown in their empirical study that lexical items are more stable in large populations and that rates of word loss are higher in small populations. Indeed, if linguistic constituents share mechanistic similarities with biological replicators (Croft, 2000; Ritt, 2004) the latter observation is exactly what one would expect as per evolutionary theory (Bromham et al. 2015: 2100).

Purely computational approaches to this problem have been, to our knowledge, primarily limited to simulations (Nettle, 1999; Wichmann, Stauffer, Schulze, & Holman, 2008). More recent advances in mathematical ecology (in particular, stochastic epidemiological dynamics; Gray, Greenhalgh, Hu, Mao, & Pan, 2011; Greenhalgh, Liang, & Mao, 2015) allow for a more analytical assessment. This paper adds to the discussion about the relationship between

This paper is distributed under a Creative Commons CC-BY-ND license. DOI:10.12775/3991-1.004 population size, linguistic stability and evolution by modifying and analyzing an established population-dynamical model of linguistic spread (Cavalli-Sforza & Feldman, 1981; Nowak, 2000; Nowak, Plotkin, & Jansen, 2000; Solé, Corominas-Murtra, & Fortuny, 2010; Wang & Minett, 2005). We focus on the dynamics of single 'linguistic items' or 'constituents' (like phonemes, n-phones, words or constructions) in finite speaker populations. After discussing the deterministic dynamics, we also analyze a stochastic version of the model, which accounts for variability in the process of constituent learning (e.g. varying density of the speaker network due to eco-linguistic factors, or varying usage of the constituent in learner-user interactions). It is shown that the general assumption that linguistic stability increases with population size only holds if variability in the learning process is kept low, and argue that the latter factor provides an interesting mechanism in language evolution.

2. Modeling linguistic spread in finite populations

2.1. Deterministic model

We study a modified version of Nowak's (2000) basic model of linguistic spread. In our version of the model, population size N is restricted to be finite. The model describes the dynamics of a structured population composed of users of a particular linguistic item i (e.g. phoneme, n-phone, word or construction) and learners that do not use it. Let U_i and L_i denote the respective sizes of the (disjoint) subpopulations and let $U_i + L_i = N$. Whenever learners and users meet, the former learn *i* at a rate λ so that they switch from class L_i to U_i . We assume λ to denote the *learning* rate, where learning of a new form is not necessarily restricted to the first years of language acquisition. Rather, we mean any interaction of individuals one of which does not yet know and use a given item. In Nowak's (2000) model, this rate λ is a function of (a) *network density*, linked to the number of communicative encounters a learner is exposed to, (b) production rate, i.e. the extent to which the item is produced, and (c) *learnability*, i.e. the probability that the item is successfully acquired when a learner is exposed to it. Learners and users die at a normalized mortality rate of 1 (so that each time unit equals one speaker generation), and dead learners and users are immediately replaced by new individuals that are added to the learner class so that population size is kept constant. In addition, users can switch back to class L_i at a rate γ when they stop using *i* ('unlearning'), for instance because they forget the item or because they abandon it in favor of a competing linguistic variant. We suggest that γ is inversely related with the *usability* of *i* in everyday speech events in

which no user-learner interactions are involved. Table 1 summarizes the model parameters.

Variable	Linguistic and cognitive interpretation		
Ν	Total size of the population of linguistic agents composed of U users and L learners		
λ	Item-specific <i>learning rate</i> in interactions; depends on <i>network connectivity</i> (linked to number of communicative encounters), <i>production rate</i> (linked to utterance frequency and ease of production), and <i>learnability</i> (linked to ease of perception)		
γ	Rate at which individuals stop using an item (in addition to speaker death; rate of 'unlearning'); inversely related to factors enhancing <i>usability</i> (e.g. ease of memorization or ease of production); assumed to be independent from learner-user interactions		
R_0	Expected number of learners that successfully learn an innovation from a single user		

Table 1. Variables in the model and how they can be interpreted

The dynamics are determined by a deterministic two-dimensional dynamical system in continuous time which models the respective growth rates of L_i to U_i . In what follows we will omit the index *i*, for the sake of simplicity, since we only focus on the dynamics of a single item (although the parallel evolution of several items clearly can be studied as well). The model equations read:

$$dL/dt = \underbrace{\overline{\lambda LU}}_{\text{learning}} + \underbrace{\widetilde{\gamma U}}_{\text{learning}} - \underbrace{\widetilde{L}}_{\text{death}} + \underbrace{\widetilde{N}}_{\text{birth}}$$
(1)

If $\gamma = 0$ and N = 1 the dynamical system reduces to the model of linguistic spread in Nowak (2000) and Solé (2011), which is equivalent with a onedimensional model of logistic growth (although the dynamics can be modeled by a single equation, e.g. only the second one in (1), we stick to the more explicit definition for the sake of clarity).

The qualitative behavior of the model can be predicted by the basic reproductive ratio R_0 which is defined as the expected number of learners that learn an item which has been innovatively introduced into the population by a single user (cf. Nowak 2000, Heffernan, Smith, & Wahl, 2005). If $R_0 > 1$ the dynamics approach a non-trivial equilibrium so that $\hat{U} = N(1 - 1/R_0)$ users know and use the item. That is, the item is stably established in the linguistic community. If, however, $R_0 < 1$ then the dynamics approach an equilibrium in which $\hat{U}_0 = 0$ users know the item. In that case, the item drops out of usage. Thus,

the basic reproductive ratio functions as a measure of the stability of a linguistic item. For that reason we treat R_0 as a measure of diachronic stability (technically, R_0 measures the stability of the equilibrium $\hat{U}_0 = 0$; if $R_0 > 1$ then \hat{U}_0 is unstable so that the population of users persists with probability 1 if any users are added to the population; if $R_0 > 1$ then \hat{U}_0 is stable so that the population of users goes extinct with probability 1).

For the present model, the basic reproductive ratio can be shown to read $R_0 = N\lambda/(1 + \gamma)$. The formula can be intuitively understood in the following way. The expected time an individual knowing the item remains in the user class is $1/(1 + \gamma)$; based on our assumption that the item is an innovation there are (approximately) *N* individuals that do not yet know the item; and each learner acquires the item at a rate of λ . Note, crucially, that since the amount of individuals which can acquire an item from a user depends on the number of learners available in the population R_0 depends on population size. Here this dependency is linear, which is an immediate reflex of the assumption that the population is homogeneously mixed so that any user can inform any learner in the population (see Section 3 for some discussion).

We are interested in the role that population size plays for the stability of a linguistic item. The basic reproductive ratio R_0 increases with N since $\lambda/(1 + \gamma) > 0$. The larger the population, the less likely is it that R_0 falls below one so that the item would inevitably drop out of usage.

In evolutionary terms, $R_0(\lambda, \gamma)$ can be interpreted as a measure the fitness of a linguistic item (Metz, Mylius, & Diekmann, 1996). Evidently, R_0 increases with λ (because $\partial R_0 / \partial \lambda = N / (1 + \gamma) > 0$) and decreases with γ (because $\partial R_0 / \partial \lambda = -N\lambda/(1 + \gamma)^2 < 0$). Thus, items with high learning rates and high usability should be selected for. That is, items are expected to evolve in such a way that they maximize ease of acquisition, production and use (probably governed by some trade-off among these factors). Moreover, the effect of optimizing λ and γ gets stronger the larger the population size N, so that items are expected to be less optimized in small populations.

2.2. Stochastic model

Things get slightly more complicated when variability in the model dynamics is considered. For instance, demographic variability could be accounted for, i.e. fluctuations due to random speaker deaths and births in addition to the deterministic model dynamics. For the class of models (1) belongs to, it has been shown that the effects of demographic variability can be neglected if population size is substantially large (Greenhalgh et al. 2015).¹ Another source of variability might be more relevant to linguistic dynamics, namely that of parametric (or environmental) variability. Here, model parameters fluctuate randomly, thus affecting the behavior of all individuals in the population at the same time. In a linguistic setting, for instance, network density of the entire speech community could vary due to eco-linguistic factors (e.g. migration or areal expansion; cf. Mufwene, 2001; Lupyan & Dale, 2010). Likewise, frequency of use of an item established in a speech community might fluctuate due to socio-linguistic or language-internal factors (e.g. morpho-syntactic or phonological restructuring, or emergence of competing variants for instance in language contact). All of these factors can be argued to have an impact on the linguistic learning process. Thus, we include a stochastic component into the model by extending the rate of transition from class L to class U, denoted by $\tilde{\lambda}$, so that $\tilde{\lambda} dt = \lambda dt + \sigma dW(t)$. Here, W(t) is a Wiener process (random noise) which accounts for fluctuation around λ , and $\sigma \ge 0$ is the variance in the 'learning environment' due to the above-mentioned factors. Thus, σ measures the magnitude of these fluctuations. We consider learning environments with low σ as more stable that those with large σ .² By replacing λdt by $\tilde{\lambda} dt$ in (1), the model becomes a system of stochastic differential equations (SDE; Allen, 2010):

$$dL = (-\lambda LU + \gamma U - L + N)dt - \sigma LUdW(t)$$

$$dU = (\lambda LU - (1 + \gamma)U)dt + \sigma LUdW(t)$$
(2)

Clearly, if there is no fluctuation ($\sigma = 0$), (2) reduces to the deterministic model (1). System (2) belongs to the class of Itô SDEs analyzed by Gray et al. (2011). Hence, we can employ the conditions for extinction and persistence derived there. By applying Theorem 4.1 in Gray et al. (2011), the basic reproductive ratio for system (2) can be shown to read

$$R_0 = \frac{\lambda N}{\underbrace{1+\gamma}_{(i)}} - \underbrace{\frac{\frac{1}{2}\sigma^2 N^2}{1+\gamma}}_{(ii)}$$
(3)

where part (i) equals the basic reproductive ratio of the deterministic system (1) and part (ii) comes from the diffusion term in the SDE (2). Theorem 5.1 in Gray et al. (2011) entails that the system leads to persistence of an item (i.e. stable and positive *U*), if $R_0 > 1$. If, on the contrary, $R_0 < 1$ and $\sigma \le \sqrt{\lambda/N}$ (Thm 4.1), or

¹ Based on Greenhalgh et al. (2015, Theorem 4.1), demographic variability only has an additional effect if population size falls below critical size $N_{\text{crit}} = 1/4 + (1 + \gamma)/\lambda$.

² Note that this notion of stability differs from the one measured by the basic reproductive ratio. While σ measures how constantly transmission of an item takes place, R_0 measures whether or not an item persists in the speaker population.

 $\sigma > \sqrt{\lambda/N}$ (Thm 4.3), then the number of users *U* approaches zero with probability 1, so that the item goes extinct.

Several observations can be made. To begin with, it is not difficult to see that $dR_0/dN > 0$ if $\sqrt{\lambda/N} > \sigma$. This means that the stability of an item increases with population size *N* as long as variability is not too high. In particular, inequality $\sigma > \sqrt{\lambda/N}$ is favored to hold (a) if learning variability σ is large or (b) if population size is high (or both). Thus, severe fluctuations promote the loss of items and impede the establishment of new items in the speaker population (Figure 1). Moreover, for fixed σ , larger population sizes can also have negative effects on the stability of linguistic items. In large populations, even mild fluctuations can yield severe reflexes, as long as they affect the entire linguistic population.



Figure 1. On the left: Itô-process simulations of diachronic developments (N = 50, $\lambda = 0.5$, $\gamma = 1.5$, U(0) = 1) in two different environments; lower variability ($\sigma = .073$, $R_0 = 7.33$, light gray), and higher variability ($\sigma = .135$, $R_0 = 0.89$, dark gray). After about 3 generations, the item exposed to higher variability in the learning environment goes extinct, as expected. On the right: R_0 as a decreasing function of σ for three different population sizes N = 50; 75; 100 ($\lambda = 0.5$, $\gamma = 1.5$ fixed). For high σ , larger populations yield lower R_0 . Computations were done in Mathematica (Wolfram Research, 2016).

What is more interesting is this: a sensitivity analysis reveals information about the relative importance of λ and γ in the optimization of R_0 in the stochastic model. For the respective directional derivatives of $R_0(\lambda, \gamma)$, we have that $\partial R_0(\lambda, \gamma)/\partial(1,0) = N/(1 + \gamma) > 0$, and that $\partial R_0(\lambda, \gamma)/\partial(0, -1) = 1/2 \cdot$ $N(2\lambda - \sigma^2 N^2) \log(1 + \gamma) > 0$, because $\sqrt{\lambda/N} > \sigma$ if the item already exists stably. Items benefit from increasing λ and decreasing γ (i.e. increasing usability), but in contrast to the former parameter, the effect of decreasing γ suffers from variability in the learning environment. For an item, to put it casually, it pays off to put more effort into improving learning rather than usability if variability is high enough. Improving factors that determine learning does always contribute to an item's success, while effects of increased usability may be vanishingly small in the presence of noise. As in the deterministic case, the effects of optimizing R_0 (i.e. the directional derivatives shown above) get stronger the larger the speaker population (Figure 2).



Figure 2. Fitness landscape defined by R_0 as a function of γ and λ for three different population sizes (N = 50; 75; 100) in the presence of learning variability ($\sigma = .073$); fitness increases linearly with λ and decreases convexly with γ . Directional slopes (effects of changing parameters) get steeper as population size increases.

3. Discussion and conclusion

By studying systems of ODEs and SDEs, we have shown that population size in general increases the stability of linguistic constituents (cf. Table 2) and thus (a) facilitates their establishment in the speaker population and (b) prevents their loss. This goes in line with Bromham et al. (2015: 2100) who show that "Polynesian languages with larger speaker-population sizes [have] higher rates of gain of new words than their smaller sister languages" and that "languages with a smaller number of speakers [have] higher rates of loss of lexemes". Our results also converge with studies that found a positive correlation between population size and the size of a language's phoneme inventory (see Nettle, 2012 for a review), and by implication phonotactic richness (Maddieson, 2013).

However, the presence of variability in the learning environment decreases stability, and the negative effects of variability get stronger, the larger population size. In the extreme case, this variability can lead to the loss of a constituent (Figure 1, left, dark gray trajectory). As a corollary of this, we can conclude that the establishment of an inventory of constituents (e.g. lexicon of words or phoneme inventory) requires a relatively stable learning environment (cf. McMahon & McMahon, 2013, p. 248). For instance, it can be argued that the small phoneme inventories found in a number of remote Polynesian languages (Trudgill, 2004) might be a reflex of migration and concomitant variability in network density. We argue that complementary to demographic variability (linked to linguistic founder effects as suggested by Atkinson 2011), environmental variability provides another interesting mechanism for explaining linguistic evolution, because it applies even if population size remains constant.³

Feature	Deterministic model	Stochastic model
Learning environment	Constant ($\sigma = 0$)	Variable ($\sigma > 0$)
Effect of <i>N</i> on stability of constituent	Stability increases with population size N	Stability increases with N if variability σ is small
Effect of <i>N</i> on evolution of learning rate λ	Adaptive effects of improving learning increase with <i>N</i>	Adaptive effects of improving learning increase with <i>N</i>
Effect of <i>N</i> on evolution of usability $\sim \gamma^{-1}$	Adaptive effects of improving usability increase with <i>N</i>	Adaptive effects of improving usability increase with <i>N</i>
Effect of variability on evolution of learning rate λ	Improving learning rate always increases fitness	Improving learning rate always increases fitness
Effect of variability on evolution of usability $\sim \gamma^{-1}$	Improving usability always increases fitness	Effects of improving usability are mitigated by variability σ

Table 2. Results and model comparison

One might wonder, what the prediction of the model, that constituent inventories are more likely to shrink in small populations actually means. Clearly, it is not plausible that small populations simply drop constituents like phonemes or lexemes, since some items obviously fulfil specific functions in the linguistic system and cannot be arbitrarily left away. Models like the ones studied in this paper cannot easily account for such details. However, one way of looking at this prediction is this: if constituents vanish (e.g. due to bad adaptation) the language must compensate for this loss, e.g. by adding more complex morpho-syntactic rules. Indeed, this is supported by Lupyan and Dale (2010) who show that small

³ Indeed, Bybee (2011) has contested demographic variability as the main explanatory link between linguistic evolution and population size.

populations sustain morphologically more complex languages.⁴ This argument contrasts the causal directionality proposed by Nettle (2012) who argues that it is the smaller number of contacts in small populations that promotes the acquisition of complex morphology (which, in turn, would allow for a reduced lexicon).

The findings also agree with Fay and Ellison (2013: 7) in the sense that increased population size enhances the optimization of properties associated with linguistic transmission. That is, evolution proceeds faster in large populations. At first sight, this may seem paradox: population size is predicted to increase the stability of an item, but at the same time population size drives linguistic optimization, where an item is effectively replaced by a more successful version of itself. Note, crucially, that the more optimized variant is less likely to get lost.

The analysis of the stochastic model has revealed that constituents always benefit from optimizing factors related to learning while advantages gained from optimizing factors related to usability can be lost due to random fluctuations in the learning environment. Based on this, it can be expected that items are relatively more optimized for being learned easily rather than for ease of use outside of the learning context. This accords with studies that propose a strong connection between (diachronic) stability and ease of acquisition (e.g. Monaghan 2014). It is less compatible with studies stressing the importance of usability and ease of production (i.e. speaker-over-listener dominance) in linguistic transmission (Bybee, 2010; Fay & Ellison, 2013).⁵

Finally, a more technical caveat is in order. The model builds on the assumption that the learning process depends on a mass-action law (i.e., interactions are proportional with the product of the number of learners and users). It has been pointed out (de Jong, Diekmann, & Heesterbeek, 1996), that this assumption does not hold in large populations in realistic ecological scenarios. Consequently, the effect of population size on the basic reproductive ratio is probably overestimated as populations become larger. Accounting for these issues

⁴ Note that this observation does not directly follow from the present analysis but rather represents a tentative hypothesis which is compatible with our results. It would be interesting, however, to study a model which includes the possibility of combining items (perhaps similar to the approach adopted by Nowak et al. 2000) to account for complexity. We would like to thank an anonymous reviewer for pointing this out.

⁵ This observation, however, might be grounded in the abstract and simplified way in which learning and using constituents is built into the model. Arguably, the rough distinction between factors relevant to learning interactions and those not associated with interactions is very simplistic and must be refined in order to capture learnability and usability more accurately.

eventually requires the implementation of a more complicated network structure.⁶ The observations made in this contribution, nevertheless, do not contradict with results from network epidemiology. In large networks, the invasion threshold vanishes under the assumption of a more realistic network structure (small world; scale free). As a consequence of the presence of super spreaders, items can spread easily through large populations (Barabási, 2016). The effects of fluctuations during the learning process in more realistic networks, though, is yet to be looked at more closely.

Acknowledgements

I would like to thank T. Mark Ellison for raising my interest into the discussed phenomenon, and Klaus Hofmann, Niki Ritt as well as three anonymous reviewers for many helpful comments on this contribution. This research was funded by FWF (Austrian Science Fund, grant No. P27592-G18).

References

- Allen, L. J. S. (2010). An introduction to stochastic processes with applications to biology. Boca Raton, Florida: CRC Press.
- Atkinson, M., Kirby, S., & Smith, K. (2015). Speaker Input Variability Does Not Explain Why Larger Populations Have Simpler Languages. *PloS one*, 10(6), e0129463. https://doi.org/10.1371/journal.pone.0129463
- Atkinson, Q. D. (2011). Phonemic diversity supports a serial founder effect model of language expansion from Africa. *Science*, 322(6027), 346–349.
- Barabási, A.-L. (2016). *Network science*. Cambridge: Cambridge University Press.
- Bromham, L., Hua, X., Fitzpatrick, T., & Greenhill, S. J. (2015). Rate of language evolution is affected by population size. *Proceedings of the National Academy of Sciences*, 112(7), 2097–2102.
- Bybee, J. (2010). *Language, Usage and Cognition*. Cambridge: Cambridge University Press.
- Bybee, J. (2011). How plausible is the hypothesis that population size and dispersal are related to phoneme inventory size? Introducing and commenting on a debate. *Linguistic Typology*, *15*, 147–153. https://doi.org/10.1515/LITY.2011.009

 $^{^{6}}$ As one of the reviewers of this contribution has rightfully pointed out, reproductive success is – at least partially – a function of network structure, which in turn determines the number of interactions. Here, the functional relationship between population size and network structure is crucial.

- Cavalli-Sforza, L. L., & Feldman, M. W. (1981). *Cultural transmission and evolution: aquantitative approach*. Princeton: Princeton University Press.
- Croft, W. (2000). *Explaining language change: An evolutionary approach*. *Longman linguistics library*. Harlow, England, New York: Longman.
- Fay, N., & Ellison, T. M. (2013). The cultural evolution of human communication systems in different sized populations: usability trumps learnability. *PloS one*, 8. https://doi.org/10.1371/journal.pone.0071781
- Gray, A., Greenhalgh, D., Hu, L., Mao, X., & Pan, J. (2011). A stochastic differential equation SIS epidemic model. SIAM Journal on Applied Mathematics, 71(3), 876–902.
- Greenhalgh, D., Liang, Y., & Mao, X. (2015). Demographic stochasticity in the SDE SIS epidemic model. *Discrete and Continuous Dynamical Systems-Series B*, 20(9), 2859–2884.
- Hay, J., & Bauer, L. (2007). Phoneme inventory size and population size. Language, 83(2), 388–400.
- Heffernan, J., Smith, R., & Wahl, L. (2005). Perspectives on the basic reproductive ratio. *Journal of The Royal Society Interface*, 2(4), 281–293. https://doi.org/10.1098/rsif.2005.0042
- Jong, M. de, Diekmann, O., & Heesterbeek, J. A. P. (1996). How does transmission of infection depend on population size? In *Epidemic models: their structure and relation to data* (pp. 84–94). Cambridge, UK: Cambridge University Press.
- Lupyan, G., & Dale, R. (2010). Language structure is partly determined by social structure. *PloS one*, 5. https://doi.org/10.1371/journal.pone.0008559
- Maddieson, I. (2013). Syllable Structure. In M. S. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Max Planck Digital Library. Retrieved from http://wals.info/chapter/12
- McMahon, A. M. S., & McMahon, R. (2013). Evolutionary linguistics. Cambridge textbooks in linguistics. Cambridge [England], New York: Cambridge University Press.
- Metz, J.A.J., Mylius, S. D., & Diekmann, O. (1996). When Does Evolution Optimize? On the Relation Between Types of Density Dependence and Evolutionarily Stable Life History Parameters. *IIASA Working Papers*, 96(004).
- Mufwene, S. S. (2001). *The ecology of language evolution*. Cambridge: Cambridge University Press.
- Nettle, D. (1999). Is the rate of linguistic change constant? *Lingua*, *108*(2), 119–136.
- Nettle, D. (2012). Social scale and structural complexity in human languages. *Phil. Trans. R. Soc. B*, 367, 1829–1836.

- Nowak, M. A. (2000). The basic reproductive ratio of a word, the maximum size of a lexicon. *Journal of theoretical biology*, 204(2), 179–189. https://doi.org/10.1006/jtbi.2000.1085
- Nowak, M. A., Plotkin, J., & Jansen, V. (2000). The evolution of syntactic communication. *Nature*, 404(6777), 495–498.
- Ritt, N. (2004). Selfish sounds and linguistic evolution: A Darwinian approach to language change. Cambridge: Cambridge University Press.
- Solé, R. V., Corominas-Murtra, B., & Fortuny, J. (2010). Diversity, competition, extinction: the ecophysics of language change. *Journal of The Royal Society Interface*, 7(53), 1647–1664. Retrieved from http://groups.lis.illinois.edu/amag/langev/paper/sole2010diversitygsc.html
- Trudgill, P. (2004). Linguistic and social typology: The Austronesian migrations and phoneme inventories. *Linguistic Typology*, 8(3), 305–320.
- Wang, W., & Minett, J. (2005). The invasion of language: Emergence, change and death. *Trends in Ecology and Evolution*, 20, 263–269.
- Wichmann, S., & Holman, E. (2009). Population size and rates of language change. *Human Biology*, 81(2), 259–274.
- Wichmann, S., Rama, T., & Holman, E. (2011). Phonological diversity, word length, and population sizes across languages: The ASJP evidence. *Linguistic Typology*, 15(2), 177–197.
- Wichmann, S., Stauffer, D., Schulze, C., & Holman, E. (2008). Do language change rates depend on population size? *Advances in Complex Systems*, 11(3), 357–369.
- Wolfram Research, I. (2016). Mathematica. Champaign, Illinois.

CAN LANGUAGES WITHOUT WRITING SYSTEMS PROVIDE NEW INSIGHTS IN LANGUAGE EVOLUTION?

CHRISTINA BEHME

christinabehme@gmail.com Department of Philosophy, Mount Saint Vincent University, Halifax, Canada

Given that (i) the first human languages presumably were simple and (ii) existing cognitive biases push language toward simplification and reduction it is perplexing that highly complex languages could evolve. Here I'll explore a potential solution for seeming paradox, that could also have implications for broader language evolution questions: the advent of writing. I hypothesize that writing allows humans to overcome the constraints of the "here and now bottleneck" and provides the foundation for linguistic innovations that feed back into spoken language. Systematically comparing languages that do and do not have writing systems could provide supporting evidence for this hypothesis. I propose that adopting this theoretical perspective could motivate important empirical work on language evolution.

1. Introduction

Language evolution theorizing faces many unique challenges, one of which is that we have no direct evidence for the earliest human languages: "Faded as soon as it is uttered, spoken language leaves no trace" (Burling 2005, 1). To circumvent this problem Burling suggests to depart from "two reasonably solid anchor points ... the behaviour of our closes primate cousins [and] modern human language" (Burling 2005, 2-3). Following this strategy immediately reveals a new puzzle. On the one hand, based on primate research (e.g. Arnold & Zuberbuehler 2008, Call & Tomasello 2007, Savage-Rumbaugh & Fields 2000) we can be rather confident that first emerging languages of our distant ancestors were fairly simple. On the other hand, modern human languages are intricately complex. Given that currently existing cognitive biases push language toward simplification and reduction, one has to wonder how complex languages could have arisen. It has been suggested that "the pressure toward reduction is normally kept in balance by the need to maintain effective communication" (Christiansen and Chater 2016 a). But this suggestion only explains why modern languages are not any simpler, not how complex languages could evolve in the first place.

This paper is distributed under a Creative Commons CC-BY-ND license. DOI:10.12775/3991-1.005

2. Language Acquisition provides clues for Language Evolution

Intuitively, adopting Ernst Haeckel's slogan 'Ontogeny recapitulates Phylogeny' might provide important insights in the evolution of a complex (adult) language from a much simpler (early child) language. And indeed, in several recent publications it was suggested that language evolution and language acquisition should not be studied in isolation but as interlinked phenomena (e.g. Christiansen and Chater 2016 a, b). For example, (Christiansen and Chater 2016a) promise to provide "an integrated framework for explaining many aspects of language structure, acquisition, processing, and evolution that have previously been treated separately". Their main argument is that languages evolved to fit cognitive constraints on how linguistic structures can be learned and how utterances can be processed in real time. In turn language has an influence on cognitive mechanisms, favouring those that are heavily used in language production and comprehension. Christiansen and Chater proposed that "gradual increases in complexity can happen relatively quickly, as indicated by the fact that children can "outperform" the adults from whom they learn language (Singleton & Newport 2004)". But, presumably, any newly emerging complexity would be subject to reduction and simplification pressures resulting in little or no net increase of complexity, unless there is a reason to maintain complexity. While the drive to communicate could provide such pressure, it has been argued convincingly that the complexity of grammar actually needed to support most daily activities of humans living in complex contemporary societies (let alone in environments our distant ancestors found themselves in) is substantially less than that exhibited by any contemporary human language (Gil 2009). Furthermore, it is unclear why certain features found in many languages (e.g. long distance dependencies) would arise in the first place: "If learners look first for local associations in blindly segmenting their language, subject to a crippling limit on short-term memory, it is unclear how long-distance dependencies could be stable in any lineage, much less universal" (Medeiros et al. 2016).

3. Writing it down

Spoken language's fleeting nature can be overcome by writing. Already Plato quotes an Egyptian myth according to which "[the invention of writing] will make the Egyptians wiser and give them better memories; for it is an elixir of memory and wisdom" (Phaedrus 274c).

Spoken language is not only fleeting but also one-dimensional (we can only speak and hear one word at a time). Yet, the world we experience is multidimensional and we routinely perceive several stimuli simultaneously. Story telling allows adding dimensions in language use but our memory limits how complex a spoken story can be. Writing is a tool that can overcome such limits. Once a word is written down it remains 'in place', allowing the reader to return to it at any time s/he wishes. Proponents of the "extended mind hypothesis" suggest that during many complex cognitive tasks "the individual brain performs some [cognitive] operations, while others are delegated to manipulations of external media" (Clark & Chalmers, 1998; see also Logan, 2007). Some commentators of Christiansen &Chater (2016a) stressed that written language could overcome some of the constraints imposed by the 'here and now bottleneck': "By the nature of texts as static visual objects, the effects of temporal constraints on information intake may be reduced or abolished. ... We acquire a portion of our vocabulary and grammar through written language, and we massively use text to communicate" (Baggioa and Vicario, 2016). Offloading part of the cognitive task to external media (the page of a book or the screen of a computer) also gives words 'permanence' and written sentences can greatly exceed the complexity of spoken sentences as this example illustrates:

> An dem Schnittpunkte von Kurfürstendamm und Kurfürstenstraße, schräg gegenüber dem »Zoologischen«, befand sich in der Mitte der siebziger Jahre noch eine große, feldeinwärts sich erstreckende Gärtnerei, deren kleines, dreifenstriges, in einem Vorgärtchen um etwa hundert Schritte zurückgelegenes Wohnhaus, trotz aller Kleinheit und Zurückgezogenheit, von der vorübergehenden Straße her sehr wohl erkannt werden konnte. Was aber sonst noch zu dem Gesamtgewese der Gärtnerei gehörte, ja die recht eigentliche Hauptsache derselben ausmachte, war durch eben dies kleine Wohnhaus wie durch eine Kulisse versteckt, und nur ein rot und grün gestrichenes Holztürmchen mit einem halb weggebrochenen Zifferblatt unter der Turmspitze (von Uhr selbst keine Rede) ließ vermuten, daß hinter dieser Kulisse noch etwas anderes verborgen sein müsse, welche Vermutung denn auch in einer von Zeit zu Zeit aufsteigenden, das Türmchen umschwärmenden Taubenschar und mehr noch in einem gelegentlichen Hundegeblaff ihre Bestätigung fand. (Fontane, 1888, 1)

The use of written language allows Fontane to paint a detailed, multilayered picture of a complex scene in just two sentences. This is possible, in part, because German has evolved into a language with complex inflection and cases system which enables the reader to keep track of which adjectives modify which subjects even over considerable distance (e.g. "... welche Vermutung denn auch in einer von Zeit zu Zeit aufsteigenden, das Türmchen umschwärmenden Taubenschar und mehr noch in einem gelegentlichen Hundegeblaff *ihre Bestätigung fand*.)

This example might demonstrate how an intricately complex grammar, once evolved, allows readers easily to follow complex stories. But, it seems unlikely that this intricately complex grammar sprang into existence instantaneously and, once more, the question arises how initial complexification arose. Plato, again, may suggest an answer: "when [words] have once been written down they are tumbled about anywhere among those who may or may not understand them ... and they have no [speaker] to protect them; and they cannot protect or defend themselves" (Phaedrus, 275c). Plato's point is crucial. During a verbal conversation the listener can ask for clarification and the speaker can provide as much detail as is needed for the successful transfer of the intended message. But the pages of a book cannot perform this task. Here only those messages can be "successful" that anticipate and answer questions of actual and potential readers. The written text has to anticipate questions, reduce ambiguity, and increase informative content. In order to accomplish this task reliably written languages needed to evolve a kind of precision that is different from that of required in spoken language. At the same time, the permanence of written language allowed for innovations inaccessible to spoken language that could provide this new kind of precision. The previously mentioned complex inflection and cases system of German is one way of solving the problem. Christiansen and Chater (2016b) stress the importance of linguistic experience for language learning and processing, and highlight reading as one important source of experience. Written texts allow for easier integration and fixation of such complex language properties as long distance dependencies, object relative clauses, and multiply embedded recursive clauses. Given that those properties add expressive power to languages they provide tangible benefits. Having a medium that overcomes the fleeting nature of spoken language could have provided an anchor point for those structures. Once they became part of written language they would have also been used increasingly in spoken language and, perhaps impacted the brains' ability to process those structures.

4. From philosophical speculation to scientific research

Virtually all languages that have been investigated by language evolution researchers had writing systems for hundreds or even thousands of years. If the invention of writing was indeed the launch-pad to any Baldwinian-type evolution, then most currently researched languages would have been affected by this change. It has been suggested that "... the earliest written documents already display the full expressive variety and grammatical complexity of modern languages" (Jackendoff 1993, 32). Additionally, many experimental designs rely heavily on written language (e.g. self-paced reading tasks, word recognition tasks, etc.). As a consequence, we may be learning more and more about the evolution of written language and making some erroneous assumptions about the earlier evolution of spoken languages.

To test the hypothesis I introduced in section 3 empirical one would need to conduct detailed comparisons of complexity parameters of languages that do and that do not have a writing system. "Ethnologue (20th edition) has data to indicate that of the currently listed 7,099 living languages, 3,866 have a developed writing system ... The remaining 3,233 are likely unwritten" (Simons & Fennig 2017). Those 'remaining languages' could provide a (albeit rapidly narrowing) window into earlier evolutionary stages of the interplay between cultural and biological language evolution. Especially language acquisition research of languages without writing systems might be a promising research avenue. Currently studies on such languages are rare (Gordon et al. In progress for Pirahã) and detailed comparative work between languages that do and do not have writing systems is nonexistent. Given that researchers are often influenced by expectations what they might find in a language (e.g. Everett, 2009) such research is especially challenging.

In part comparative research between languages that do and that do not have a writing system might have been hampered by a "longstanding linguistic axiom - that all languages ... are similar in complexity, and have been so at all times in the past" (Sampson, 2009a). Researchers paralyzed by ideological fears should remember that language evolution is not a race towards some perfect language and, therefore, more complex languages, or languages with a writing system are not 'superior' to others. Rather languages ought to be evaluated by how well they fit the communicative needs of their respective communities. From that perspective, a language that allows the complex sentences of German is not any better (or worse) than a language (like Pirahã) that does not.

Of course, it has been known for a long time that "unwritten third world grammars often contain highly sophisticated structural features" (Sampson 2009b). But it would be desirable to find out whether there are systematic differences between the complexities of languages that do and that do not have a writing system. For example, it turns out that Jackendoff was wrong about the unchanging complexity of written languages. Guy Deutscher documented that "the earliest recorded stages of Akkadian [lack] finite complement clauses ... [and that] complement clauses [are] gradually developing out simpler, non-recursive which did exist in the early records" (Sampson, 2009b). It is of course possible that cultural changes, that occurred independently of the writing system, generated novel communicative needs for which recursive structures provided a solution. For this reason one needs to compare languages systematically. If it turns out that *only* languages with a writing system have certain complex properties (e.g. long distance dependencies or multiply embedded recursive clauses), then one might argue with confidence that a writing system is a necessary condition for the evolution of those language properties. If languages with a writing system are statistically more likely to have those properties, then a writing system is not necessary (but perhaps beneficial) for their evolution. And if languages with and without a writing system are equally likely to have those properties, then a writing system is irrelevant for their evolution.

Further it would be good to compare systematically the kids of messages that are most prone to error in languages with and without a writing system. Over centuries, even for language with a writing system, writing was a privilege of the educated upper classes. Their needs for communication and entertainment determined what was written town and passed on to the next generations. Only recently has literacy been extended to most speakers of many language with writing systems. And, the recent advent of Twitter has greatly reduced the linguistic complexity of some speakers and also shown that existing written language is ill adapted to unambiguous transfer of information in very limited space. Language evolution researchers might win important insights from monitoring how this novel cultural phenomenon impacts complex language properties.

Finally, it would also be important to find out whether there are complex language properties that are only found in languages without a writing system. This proposal may seem paradoxical since all languages evolved from languages without a writing system. However, because language evolution is not 'aiming' at some ideal end-sate, complexity does not always increase. In fact, Plato had already suggested that the invention of writing might have an adverse effect: " [the letters] will create forgetfulness in the learners' souls, because they will not use their memories [but trust the written words]" (Phaedrus 275a). Because a language without writing system places a greater burden on working memory it may have evolved specific cognitive tools that are no longer needed when writing allows one to 'extend' one's mind and 'offload' part of the cognitive task. Again, only detailed empirical research can (dis)confirm these speculations.

Overall, I believe that a systematic comparison between languages with and without writing systems can provide new perspectives on language evolution research and, perhaps, help answering the question "Why did complex languages evolve in spite of cognitive biases pushing towards simplification and reduction.

References

- Arnold, K. & Zuberbuehler, K. (2008) Meaningful call combinations in a nonhuman primate. *Current Biology*. 18, 5, R202-R203.
- Burling, R. (2005) The talking ape. Oxford: Oxford University Press.
- Call, J. & Tomasello, M. (2007). The gestural communication of apes and monkeys. Mahwah, NJ: Lawrence Erlenbaum.
- Christiansen, M. H., & Chater, N. (2016a). The Now-or-Never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*, 39
- Christiansen, M. H., & Chater, N. (2016b). *Creating Language*. Cambridge, MA: MIT Press.
- Clark, A. & Chalmers, D. (1998). The Extended Mind. Analysis, 58, 1, 7-19.

- Everett, D. (2009). An interview with Dan Everett. In: G. Sampson, D. Gill, and P. Trudgill (Eds.) *Language Complexity as an Evolving Variable*. (pp. 213-229). Oxford: Oxford University Press.
- Fontane, T. (1888). Irrungen Wirrungen On-line: http://gutenberg.spiegel.de/ buch/irrungen-wirrungen-4457/1 accessed 14/01/2018.
- Gil, D. (2009) How much grammar does it take to sail a boat? In: G. Sampson,D. Gill, and P. Trudgill (Eds.) *Language complexity as an evolving variable*.pp. 19–33. Oxford: Oxford University Press.
- Giosuè B. & Vicario, C. M. (2016). Language processing is not a race against time *Behavioral and Brain Sciences*, 39.
- Gordon, P., Kirby, E., Tang, J., Zaleznik, E., Maldonado, D., Orr, A., Gazman, Z., Truong, S., Cheng, R., Shi, W., Wang, Y., Madden, S., & Everett, D. (In Progress). Pirahã Motherese.
- Jackendoff, R. (1993). *Patterns in the mind: language and human nature*. New York: Basic Books.
- Logan, R. K. (2007). *The extended mind: The emergence of language, the human mind and culture*. Toronto: University of Toronto Press.
- Medeiros, D., Piattelli-Palmarini, M. & T. G. Bever (2016). Many important language universals are not reducible to processing or cognition. *Behavioral* and Brain Sciences, 39.
- Plato (1937). Phaedrus. In: Jowett, B. (Translator) *The dialogues of Plato*. pp. 233-284. New York: Random House.
- Sampson, G. (2009a). Preface. In: G. Sampson, D. Gill, and P. Trudgill (Eds.) Language complexity as an evolving variable. pp. vii-viii. Oxford: Oxford University Press.
- Sampson, G. (2009b). A linguistic axiom challenged. In: G. Sampson, D. Gill, and P. Trudgill (Eds.) *Language complexity as an evolving variable*. pp. 1-18. Oxford: Oxford University Press.
- Savage-Rumbaugh, E.S. & Fields, W.M. (2000) Linguistic, Cultural and Cognitive Capabilities of Bonobos (Pan paniscus). *Culture & Psychology 6(2)*, 131-153.
- Simons, G. F. & Fennig, C.D. (Eds.). 2017. Ethnologue: Languages of the World, Twentieth edition. Dallas, Texas: SIL International. Online version: http://www.ethnologue.com. accessed 14/01/2018.

LANGUAGE FAMILY TREES REFLECT GEOGRAPHY AND DEMOGRAPHY BEYOND NEUTRAL DRIFT

CHRISTIAN BENTZ*1,2, DAN DEDIU3, ANNEMARIE VERKERK4, and GERHARD JÄGER1,2

*Corresponding Author: chris@christianbentz.de
¹Department of General Linguistics, University of Tübingen, Tübingen, Germany
²DFG Center for Advanced Studies, University of Tübingen, Tübingen, Germany
³Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
⁴Max Planck Institute for the Science of Human History, Jena, Germany

1. Introduction

Languages constantly evolve. Language families unfold in time and space, and span the globe. A core question of evolutionary linguistics is whether this remarkable diversity is the outcome of mere neutral drift, or if further pressures relating to geography, climate, population size, and human biology guide diversification. We apply phylogenetic signal analyses to measure the reflection of longitudes, latitudes, altitudes and population sizes on language family trees. In particular, we evaluate the evidence that language family tree structure and external factors have evolved completely independently (phylogenetic signal of zero), if they have evolved by neutral drift (phylogenetic signal of one), or whether there are further adaptive and non-adaptive pressures at play (phylogenetic signal between zero and one or bigger than one).

2. Methods

2.1. Tree samples

We collect phylogenetic trees from three major sources: 1) A database harnessing openly available tree topologies (Dediu, in press). 2) Bayesian posterior trees supplied by authors of recent phylogenetic studies for overall eight language families. 3) Trees derived via the maximum likelihood (ML) method applied to ASJP word lists (Wichmann et al., 2013). Overall, we arrive at a sample of 52588 trees representing 54 language families.

2.2. Geographic and demographic information

Latitude and longitude information per language is taken from Glottolog (Hammarström, Forkel, Haspelmath, & Bank, 2016). Altitudes are estimated using the

This paper is distributed under a Creative Commons CC-BY-ND license. DOI:10.12775/3991-1.006 *Google Maps Elevation API*¹. Population size data is taken from the Ethnologue (Lewis, Simons, & Fenning, 2013). We arrive at a sample of 6834 languages (unique ISO-639-3 and glottocodes) for which latitude, longitude, altitude, and population size is available.

2.3. Phylogenetic signal metrics

In a linguistic context, phylogenetic signal can be conceptualized as the extent to which languages close to each other on a phylogenetic tree also resemble each other in their population-level "external" traits. There is a range of methods to estimate phylogenetic signal (Münkemüller et al., 2012; Blomberg & Garland, 2002). We focus here on two metrics in particular: *Blomberg's K* (Blomberg, Garland Jr, Ives, & Crespi, 2003), and *Pagel's \lambda* (Pagel, 1997).

3. Results and Discussion

We investigate 4 external factors by both signal metrics, K and λ , for all three sources of trees. For these 24 combinations, geographic factors and population size have mean phylogenetic signals significantly higher than zero (Wilcoxon test: p < 0.05). The mean values across the 12 combinations per signal metric range from $\bar{\lambda} = 0.19$ (SD = 0.26) to $\bar{\lambda} = 1.001$ (SD = 0.071), and from $\bar{K} = 0.35$ (SD = 0.15) to $\bar{K} = 4.32$ (SD = 4.54) respectively. These are also generally significantly different from one (p < 0.05). Analyses for 128 combinations by signal metric, external factor and language family yield qualitatively similar results.

Hence, phylogenetic signals are generally stronger than expected under the null hypothesis of completely independent evolution (i.e. $\bar{K} = 0$ and $\bar{\lambda} = 0$). Moreover, the prediction of neutral drift (i.e. $\bar{K} = 1$ and $\bar{\lambda} = 1$) is mostly not confirmed. This suggests that there are further adaptive and non-adaptive pressures at play when language families evolve and diversify. Based on systematic differences between language families we argue that these include convergent evolution, niche occupancy, heterogeneous rate drift and lateral transfer of lexical and structural material. Finally, there is preliminary evidence that language families expand more along longitudes than latitudes, thus confirming the hypothesized east-west tendency of human migrations (Diamond, 1999; Güldemann & Hammarström, in print).

Acknowledgements

CB and GJ were funded by the German Research Foundation (DFG FOR 2237: Project "Words, Bones, Genes, Tools: Tracking Linguistic, Cultural, and Biological Trajectories of the Human Past") and by the ERC Advanced Grant 324246 EVOLAEMP.

¹https://developers.google.com/maps/documentation/elevation

References

- Blomberg, S. P., & Garland, T. (2002). Tempo and mode in evolution: phylogenetic inertia, adaptation and comparative methods. *Journal of Evolutionary Biology*, 15(6), 899–910.
- Blomberg, S. P., Garland Jr, T., Ives, A. R., & Crespi, B. (2003). Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution*, 57(4), 717–745.
- Dediu, D. (in press). Making Genealogical Language Classifications Available for Phylogenetic Analysis: Newick Trees, Unified Identifiers, and Branch Length. *Language Dynamics and Change*.
- Diamond, J. M. (1999). *Guns, germs and steel: the fates of human societies*. New York/ London: W. W. Norton.
- Güldemann, T., & Hammarström, H. (in print). Geographical axis effects in largescale linguistic distributions. In M. Crevels & P. Muysken (Eds.), *Language dispersal, diversification and contact*. Oxford: Oxford University Press.
- Hammarström, H., Forkel, R., Haspelmath, M., & Bank, S. (Eds.). (2016). Glottolog 2.7. Jena: Max Planck Institute for the Science of Human History.
- Lewis, M. P., Simons, G. F., & Fenning, C. D. (Eds.). (2013). *Ethnologue: Languages of the world* (17th ed.). Dallas, Texas: SIL International.
- Münkemüller, T., Lavergne, S., Bzeznik, B., Dray, S., Jombart, T., Schiffers, K., & Thuiller, W. (2012). How to measure and test phylogenetic signal. *Methods in Ecology and Evolution*, 3(4), 743–756.
- Pagel, M. (1997). Inferring evolutionary processes from phylogenies. Zoologica Scripta, 26(4), 331–348.
- Wichmann, S., Müller, A., Wett, A., Velupillai, V., Bischoffberger, J., Brown, C. H., Holman, E. W., Sauppe, S., Molochieva, Z., Brown, P., Hammarström, H., Belyaev, O., List, J.-M., Bakker, D., Egorov, D., Urban, M., Mailhammer, R., Carrizo, A., Dryer, M. S., Korovina, E., Beck, D., Geyer, H., Epps, P., Grant, A., & Valenzuela, P. (2013). *The ASJP database (version 16)*.

DO NON-NATIVE SPEAKERS CREATE A PRESSURE TOWARDS SIMPLIFICATION? CORPUS EVIDENCE

ALEKSANDRS BERDICEVSKIS*1

*Corresponding Author: aleksandrs.berdicevskis@lingfil.uu.se ¹Department of Linguistics and Philology, Uppsala University, Uppsala, Sweden

An influential paradigm within evolutionary linguistics is that languages change in response to socioecological pressures. Language complexity is a common parameter to test for such adaptation (Beckner et al., 2009: 12; Lupyan & Dale, 2010). Strong claims have been made about the evolution of complexity. e.g. that large proportion of non-native speakers in a population facilitates morphological simplification (Trudgill, 2011). While there exists evidence in favour of this claim (Bentz & Winter, 2013; Szmrecsanyi & Kortmann, 2009; Bentz & Berdicevskis, 2016), it still rests on several assumptions that have not been rigorously tested.

One such assumption is that the linguistic production of non-native speakers tends to be simpler than that of native speakers, thus creating a pressure towards simplification. Some quantitative comparisons of the complexity of L1 and L2 production have been made (Brezina & Pallotti, 2016 and references therein), but no studies involved large corpora of natural written production.

To address this issue, I create large corpora of native and non-native English, French, Italian and Spanish by using data from WordReference forums. At these forums, users have to indicate their native language in their profiles. For each of the four languages, a forum exists where the rules permit discussions solely in this language. I download the content of these forums, noting for every post the nickname of its author and whether the author is a native speaker of the language the post is written in. The resulting corpus sizes are reported in Table 1.

As a proxy of complexity I use lexical diversity (LD), operationalized in three different ways: type-token ratio (TTR) and related, but more sophisticated measures called HD-D and MTLD. TTR has been criticized for a number of shortcomings (Jarvis, 2002), but has an advantage of being easily interpretable. HD-D and especially MTLD are claimed to be more robust and less sensitive to

This paper is distributed under a Creative Commons CC-BY-ND license. DOI:10.12775/3991-1.007 text size measures (McCarthy & Jarvis 2010). A comparison of some measures of morphological complexity (Bentz et al. 2016) suggests that TTR is doing reasonably well. Bentz et al. (2016) worked with parallel texts, which cannot be done in my situation, but I perform all comparisons on texts of equal lengths.

	Italian	French	Spanish	English
L1	3.5	6.6	22.4	49.8
L2	1	3.7	5.5	38.0

Table 1. Number of tokens in every corpus (in millions).

I define three thresholds: 200, 300 and 400 tokens. LD measures may be unreliable at lower thresholds (Koizumi & In'nami 2012), while higher thresholds yield too few datapoints. For every threshold n, the following procedure is repeated: posts shorter than n are discarded, for every other post, all three measures are calculated using the first n tokens. I fit then a mixed-effect regression model with an LD measure as the response variable, Speaker type (L1 vs. L2) as a main effect and Author as a random effect and perform a likelihood-ratio test against a null model without the Speaker type predictor.

Speaker type is a significant predictor for English (TTR and HD-D; all thresholds) and French (MTLD; threshold 200). For English, both TTR and HD-D show that L2 production is less complex, for French, MTLD shows that it is more complex. In all cases, the slopes are small, but not negligible. All other combinations of language, measure and threshold do not give significant results, but the observed differences suggest that Italian behaves like English, while Spanish behaves like French.

In the talk I review these results and potential reasons for differences between English and French. I also discuss implications for typological theories outlined in the first paragraph (can it be that L2 speakers create a pressure towards simplification in some cases, but not others?). Finally, I turn to methodological issues of measuring complexity of natural written production using unannotated corpora.

Speaking of methodological issues, it should be noted that there are several potential confounds. First, the results can be affected by orthographical variation. Second, different L1 backgrounds of L2 speakers may play a role. Factoring these parameters into analysis and including other types of measures than LD-based ones are natural further steps.

References

- Beckner, C., Blythe, R., Bybee, J., Christiansen, M., Croft, B., Ellis, N.C., Holland, J., Ke, J., Larsen-Freeman. D., & Schoenemann, T. (2009). Language is a complex adaptive system: Position paper. *Language learning* 59(s1), 1–26.
- Bentz, C.; Ruzsics, T.; Koplenig, A & Samaržić, T. (2016). A comparison between morphological complexity measures: typological data vs. language corpora. In Proceedings of the Workshop on Computational Linguistics for Linguistic Complexity (CL4LC), 26th International Conference on Computational Linguistics, Osaka, Japan, 11 December 2016.
- Bentz, C., & Berdicevskis, A. (2016). Learning pressures reduce morphological complexity: Linking corpus, computational and experimental evidence. In Proceedings of the Workshop on Computational Linguistics for Linguistic Complexity (CL4LC), 26th International Conference on Computational Linguistics, Osaka, Japan, 11 December 2016.
- Bentz, C., & Winter, B. (2013). Languages with more second language learners tend to lose nominal case. *Language Dynamics and Change 3(1)*, 1–27.
- Brezina, V., & Pallotti, G. (2016). Morphological complexity in written L2 texts. Second Language Research, 0267658316643125.
- Jarvis, S. (2002). Short texts, best-fitting curves and new measures of lexical diversity. *Language Testing*, 19(1), 57-84.
- Koizumi, R., & In'nami, Y. (2012). Effects of text length on lexical diversity measures: Using short texts with less than 200 tokens. *System*, 40(4), 554-564.
- Lupyan, G., & Dale, R. (2010). Language structure is partly determined by social structure. *PLoS ONE*, 5 (1), e8559.
- McCarthy, P. M., & Jarvis, S. (2010). MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior* research methods, 42(2), 381-392.
- Szmrecsanyi, B., & Kortmann, B. (2009). The morphosyntax of varieties of English worldwide: A quantitative perspective. *Lingua 119*(11), 1643–1663.
- Trudgill, P. (2011). Sociolinguistic typology: social determinants of linguistic complexity. Oxford: Oxford University Press.

A CRITICAL PERIOD FOR THE EVOLUTION OF LANGUAGE-READINESS: CLARIFYING THE GLOBULARIZATION HYPOTHESIS

CEDRIC BOECKX*1,2,3

^{*}cedric.boeckx@ub.edu ¹ICREA, Barcelona, Spain ²Universitat de Barcelona, Barcelona, Spain ³Universitat de Barcelona Institute of Complex Systems, Barcelona, Spain

If, as is often said, the mind is what the brain does, the evolution of our species' neuroanatomy ought to play an important role in accounting for the emergence of cognitive modernity and its most salient characteristic: our full-fledged language capacity. Traits like brain size or hemispheric lateralization have long figured prominently as factors that made the modern human brain special, but comparative research has cast doubt on explanations based exclusively or primarily on these traits. By contrast, the truly sapiens-specific brain growth trajectory that has been argued to give rise to a globular skull — the 'globularization phase' as per Hublin, Neubauer, and Gunz (2015) — remains understudied, though it constitutes a more robustly species-specific trait than better-studied neurological parameters (Boeckx, 2013). Indeed, according to the Globularization hypothesis (Boeckx, 2017), this specific brain growth trajectory played an important role in making the human brain fully language-ready.

Here I would like to clarify the content of the Globularization hypothesis, and in so doing adduce additional evidence in support of it. To begin with, and contrary to previous research on this topic, the emphasis should be on the growth curve, not the ultimate craniofacial shape. The latter is the result of several factors, early brain growth being one of them. It is, in fact, possible to identify situations where globular craniofacial shape is not accompanied by our species-brain growth trajectory (with cognitive deviance as outcome; e.g., Down syndrome). Accordingly, the search for the molecular basis of the globularization phase should not be focused on osteogenic factors (contra Boeckx and Benítez-Burraco (2014)), but rather on changes primarily affecting brain growth.

This talk will rely on a detailed analysis of the archaic human genomes currently available to characterize two sets of candidate genes harboring potentially relevant mutations. The first set consists of candidate genes for either microcephaly or macrocephaly. Based on their typical expression patterns, these genes

This paper is distributed under a Creative Commons CC-BY-ND license. DOI:10.12775/3991-1.008 may have been important for early brain growth trajectory changes. A second set of genes, implicated in synaptic plasticity, appears to reflect the need to accommodate brain growth changes at the connectivity level, postnatally. Thus, the globularization phase may be best understood as consisting of two stages: a first stage leading to accelerated brain growth around birth (thereby affecting primarily late developing regions like the cerebellum), and a second, postnatal stage affecting connectivity across brain regions. Interestingly, these two stages correspond to the two critical periods for the development of autism spectrum disorders (Parikshak et al., 2013), which suggests that a disregulation of changes underlying globularization may be at the heart of some autistic traits.

The globularization phase, once broken down into two stages, makes clear that some functional consequences of early brain growth trajectory changes may manifest themselves only much later in time. Thus, the distance across levels of analysis, from the molecular to the cognitive/behavioral, is not only to be measured in terms of space (Fisher, 2015), but also in terms of time. In addition, the functional consequences of globularization can only be understood in the context of other changes with which globularization stands in a 'feedback loop' relation: (i) the emergence of a cortical vocal learning circuit (Jarvis, 2004; Fitch, 2010), (ii) neurobiological changes leading to increased cooperation (perhaps best understood in terms of self-domestication (Theofanopoulou et al., 2017)), and (iii) contextual factors leading to increased cultural evolution.

Acknowledgements

I acknowledge the financial support from the Spanish Ministry of Economy and Competitiveness (grant FFI2016-78034-C2-1-P), a Marie Curie International Reintegration Grant from the European Union (PIRG-GA-2009-256413), research funds from the Fundació Bosch i Gimpera, and from the Generalitat de Catalunya (2014-SGR-200).

References

- Boeckx, C. (2013). Biolinguistics: forays into human cognitive biology. J. Anthropol. Sci, 91, 63–89.
- Boeckx, C. (2017). The language-ready head: Evolutionary considerations. *Psychonomic bulletin & review*, 24(1), 194–199.
- Boeckx, C., & Benítez-Burraco, A. (2014). The shape of the human languageready brain. *Frontiers in psychology*, 5.
- Fisher, S. E. (2015). Translating the genome in human neuroscience. In *The future* of the brain: Essays by the world's leading neuroscientists (pp. 149–159). Princeton University Press.

Fitch, W. T. (2010). The evolution of language. Cambridge University Press.

Hublin, J.-J., Neubauer, S., & Gunz, P. (2015). Brain ontogeny and life history in pleistocene hominins. *Phil. Trans. R. Soc. B*, *370*(1663), 20140062.

- Jarvis, E. D. (2004). Learned birdsong and the neurobiology of human language. Annals of the New York Academy of Sciences, 1016(1), 749–777.
- Parikshak, N. N., Luo, R., Zhang, A., Won, H., Lowe, J. K., Chandran, V., Horvath, S., & Geschwind, D. H. (2013). Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. *Cell*, 155(5), 1008–1021.
- Theofanopoulou, C., Gastaldon, S., ORourke, T., Samuels, B. D., Messner, A., Martins, P. T., Delogu, F., Alamri, S., & Boeckx, C. (2017). Selfdomestication in homo sapiens: Insights from comparative genomics. *PloS* one, 12(10), e0185306.

MODELLING THE EFFECT OF ICONICITY ON ITERATED (CROSS-SITUATIONAL) LEARNING

Bart de Boer1

bart@ai.vub.ac.be ¹AI-lab, Vrije Universiteit Brussel, Brussels, Belgium

Iconicity has been shown to have an important influence on both the learnability of lexical items (Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015) and on the way in which systems of signals emerge in cultural evolution (Little, Eryılmaz, & de Boer, 2017; Roberts, Lewandowski, & Galantucci, 2015). However, the effect of iconicity is different under different circumstances, sometimes facilitating learning, sometimes hindering it (reviewed in Dingemanse et al. 2015). Iconicity may help early acquisition and emergence of language, but when a communication system grows, iconicity may actually become problematic and arbitrary form-meaning mappings may be more advantageous. Thus, iconicity may have played a different role in early evolution of language than it plays today.

Because of the different effects of iconicity under different circumstances (which may not all be controllable in experiments or observational studies) and because its role may have changed during biological evolution, here an agentbased computer model is used to investigate the effect of iconicity on cultural evolution of lexical systems. The model is based on cross-situational learning, where two agents communicate about a situation in which multiple meanings are present; one agent signals a meaning; the other needs to learn which signal corresponds to which meaning. The learning agent needs to recover the correct associations between signals and meanings. The original contribution of this model is that meanings and signals may both trigger identical sensory features, and hence can be iconic. When an agent determines a signal's meaning, it not only uses the learned associations, but also the (innately determined) similarity between the signal and the meanings present. Both signals and meanings are represented by 10 continuous features with a range of [-4,4]. They share 5 of these features.

This paper is distributed under a Creative Commons BY-NC-ND license. DOI:10.12775/3991-1.009

Preliminary results of applying the model to transmission chains (in which there is one teacher and one learner, and the learner becomes the teacher in the next generation) are given in Table 1. Success of communication at the 10^{th} generation (convergence takes place rapidly, so that it is nearly always complete at this point) was observed for 1) iconic and non-iconic transmission, where 2) the initial agent was a perfect communicator (i.e. one meaning corresponds to exactly one signal and vice versa) or was initialized randomly, where 3) reproduction made use of sampling or of a-posteriory maximization (i.e. formmeaning mappings were reproduced with the probability by which they were observed, or each meaning was coupled with the most frequently observed signal), and where 4) the probability of each meaning occurring was uniform, or followed a Zipf distribution (i.e. some meanings were much more frequent than others). Lexicons consisted of 20 words, were trained for 10 000 interactions, and contexts consisted of 5 meanings. All runs were repeated 50 times with different random seeds. The parameters were set such that learning was neither too complex nor too simple

It can be observed that maximizers work much better than samplers (as was expected from Smith & Kirby, 2008), that meanings with a Zipf distribution generally fare somewhat better than uniformly distributed meanings and that maximizers manage to transmit systems quite well, and even to create them from scratch (i.e. in the case of random initializations). However, the effect of iconicity in maximizers appears very marginal, if there is an effect at all. There is a clear effect in samplers, but they do worse than maximizers in all cases.

It is perhaps logical that in transmission chains with one teacher and learner, iconicity does not have a strong influence on success in the case of maximizers: success depends on the fact that the teacher has only one signal associated with each meaning, and the learner needs to learn (and later reproduce) this. Thus, iconicity is only a distraction. However, in a social coordination setting, where multiple agents negotiate a signaling system from scratch, iconicity may play an important role in helping the agents converge to the same signal for the same meaning. Investigating this hypothesis is the next step in this modeling effort, as well as investigating what happens to the influence of iconicity when the lexicon grows and becomes conventionalized. It should also be investigated whether there are *any* situations in which samplers do better than maximizers.

Table 1. Communicative success in the tenth generation for different configurations of the model. Numbers represent: median^{3rd} quartile

Sampler		Maximizer	
uniform	Zipf	uniform	Zipf
$0.05_{0.05}^{0.05}$	$0.17^{0.17}_{0.17}$	111	111
$0.39_{0.33}^{0.47}$	$0.36_{0.31}^{0.39}$	1 ¹ ₁	$0.97^{1.00}_{0.95}$
	Sam uniform 0.05 ^{0.05} 0.39 ^{0.47} 0.39 ^{0.47}	$\begin{tabular}{ c c c c c } \hline Sampler & $Zipf$ \\ \hline $0.05^{0.05}_{0.05}$ & $0.17^{0.17}_{0.17}$ \\ \hline $0.39^{0.47}_{0.33}$ & $0.36^{0.39}_{0.31}$ \\ \hline \end{tabular}$	$\begin{tabular}{c c c c c c c c c c c c c c c c c c c $

Perfect Initialization

	Random Initialization			
not iconic	$0.05_{0.05}^{0.05}$	$0.17^{0.17}_{0.17}$	$0.70^{0.75}_{0.70}$	$0.86_{0.83}^{0.90}$
iconic	$0.41^{0.48}_{0.34}$	$0.35_{0.29}^{0.38}$	$0.75^{0.80}_{0.70}$	$0.85_{0.83}^{0.89}$

References

- Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, Iconicity, and Systematicity in Language. *Trends* in Cognitive Sciences, 19(10), 603–615.
- Little, H., Eryılmaz, K., & de Boer, B. (2017). Signal dimensionality and the emergence of combinatorial structure. *Cognition*, *168*(1), 1–15.
- Roberts, G., Lewandowski, J., & Galantucci, B. (2015). How communication changes when we cannot mime the world: Experimental evidence for the effect of iconicity on combinatoriality. *Cognition*, *141*, 52–66.
- Smith, K., & Kirby, S. (2008). Cultural evolution: implications for understanding the human language faculty and its evolution. *Philosophical Transactions of the Royal Society of London B*, 363, 3591–3603.

50 GESTURES REFLECTED – TRACING GESTURAL DEVELOPMENT FROM A CHILD'S EEG SIGNAL

MONIKA BORUTA*¹²³, JOANNA DRESZER²³, MONIKA LEWANDOWSKA²³, NATALIA PAWLACZYK²³⁴, MAGDALENA SCHMIDT²³, MAGDALENA KMIECIK²³, MONIKA GRZYBOWSKA²³, AGNIESZKA IGNACZEWSKA²³, WŁODZISŁAW DUCH²³

*Corresponding Author: monika.feba@vp.pl

 Centre for Language Evolution Studies, Nicolaus Copernicus University, Torun, Poland
 Interdisciplinary Center for Modern Technologies, Nicolaus Copernicus University, Torun, Poland
 Nicolaus Copernicus University, Torun, Poland
 Warsaw University, Warsaw, Poland

1. Introduction

The research on the way gesture and language intertwine in a child's struggle for speech is still unfolding. There are many studies concerning acquisition and use of gestures by children at an early stage of their life (e.g. Capirci et al. 1998, Liebal et al. 2009, Kidd and Holler 2009). Infants and babies incorporate gestures in their expression from early stages of their lives. The use of nonverbal signals augments a child's message and lends him or her possibilities for social contact and development (Tsao et al. 2004). Gestural expression is also a way in which the child manifests his or her temperament (Rothbart et al. 1992).

2. Theoretical Background

It has been established that both spoken and sign language development can be assessed by means of early gesture observation. For example, children who tend to point and name objects perform better in two-word stage production (e.g. Iverson and Goldin-Meadow 2005). Interestingly, in later development (3–5 years of age), children who utilize gestures along verbal narration tend to resolve lexical ambiguities better than those who do not incorporate gestures in their reasoning. An example comes from Kidd and Holler's study, where they show that in 4-year-olds polysemous expressions are explained by gestures rather than speech (2009).

Gestural expression were shown to be useful in logical thinking and problem solving (as in mathematical equations) – schoolchildren who gesturally trace elements of equations provide better answers and can faster explain the task than children who do not use gestures (Goldin-Meadow et al. 2009). These examples show the importance of nonverbal communication throughout early linguistic development of the child – when the child copies parents' behaviours (e.g. Özçalışkan and Dimitrova 2013), and later, when the child is ready to use gestural expressions to facilitate speech production and understanding of new phenomena.

This paper is distributed under a Creative Commons CC-BY-ND license. DOI:10.12775/3991-1.010 The study, has been conducted as a spinoff of the Grant Project addressing phonemic awareness and hearing in newborns and infants. The main project has been carried out in the Baby Lab of the Centre for Modern Interdisciplinary Technologies, Nicolaus Copernicus University, Torun, Poland. During the phonematic training, held in the facility, children had several meetings with a native speaker of French, who was interacting with them (reading stories, talking). The children were recorded throughout seven subsequent sessions, each lasting about 30 minutes. The recordings present the native speaker-child interactions. After preparations, the children underwent a repeated recorded EEG and Eye Tracking sessions the aim of which was to check whether they react to a stimulus containing sounds of French language.

After assessing the whole material, we elicited 10 children (8–12 months of age) whose videos met the requirements for being analysed in ELAN (angle, visibility, audibility, length) and lasted in total c.a. 40 hours. Additionally, to participate in the research, parents were obliged to fill in Mary Rothbart's Infant Behaviour Questionnaire at home, which enabled us to assess the children's temperament. Our hypotheses are encapsulated in three research questions:

- 1. Is the difference in the children's temperament visible in their gestural and motor behaviour? Bates et al.'s (1989) suggest that individual differences in gestural expression may stem from various temperaments of children. We expect to complement the research with that piece of information.
- 2. Are there temperament-dependent behavioural and nonverbal exploration patterns that a child manifests meeting-to-meeting as he or she is getting familiar with the new environment? Kuhn et al. (2014) show that gestures are a precursor of language development and executive functions (EF). We expect that the more the child uses gestures (depending on their temperament), the sooner he or she will show his or her preferences (EF) for the items and places in the room.
- 3. Do changes in a child's EEG signal (Mu signal) correspond to his or her gestural behaviour? Marshall and Meltzoff (2011) point out that a child's EEG reacts to action observation and action execution. We would like to see if the changes are visible month-to-month as children's attention develops and changes as they interact with live stimuli. We expect to see differences in the Mu signal in children presenting different temperaments, and would like to show that it is possible to see an increased/lowered gesture use in a child's EEG signal.

References:
- 5Bates, E., Thal, D., Whitesell, K., Fenson, L., & Oakes, L. (1989). Integrating language and gesture in infancy. *Developmental Psychology*. 25(6): 1004–1019
- Capirci, O., Montanari, S. and Volterra, V. (1998). Gestures, signs, and words in early language development. *New Directions for Child and Adolescent Development*. (79): 45–60.
- Goldin-Meadow, S., Cook, S.W., & Mitchell, Z. (2009). Gesture gives children new ideas about math. *Psychological Science*. (20): 267–271. 10.1111/j.1467-9280.2009.02297.
- Iverson, J.M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science*. (16): 368–71. 10.1111/j.0956-7976.2005.01542.x.
- Kidd, E., & Holler, J. (2009). Children's use of gesture to resolve lexical ambiguity. *Developmental Science*. (12): 903–913.
- Kuhn, L.J., Willoughby, M.T., Wilbourn, M.P., Vernon-Feagans, L., Blair, C.B., & the Family Life Project Key Investigators. (2014). Early Communicative Gestures Prospectively Predict Language Development and Executive Function in Early Childhood. *Child Development*. 85(5): 1898–1914. 10.1111/cdev.12249.
- Liebal, K., Behne, T., Carpenter, M., & Tomasello, M. (2009). Infants use shared experience to interpret pointing gestures. *Developmental Science*. (12): 264–271.
- Marshall, J.P., & Meltzoff, A. (2011). Neural Mirroring Systems: Exploring the EEG Mu Rhythm in Human Infancy. *Developmental cognitive neuroscience*. (1): 110–23. 10.1016/j.dcn.2010.09.001.
- Özçalıskan, Ş., & Dimitrova, N. (2013). How gesture input provides a helping hand to language development. *Seminars in Speech and Language*. 34(4): 155–164.
- Rothbart, M.K., Ziaie, H., & O'Boyle, C.G. (1992). Self-regulation and emotion in infancy. In N. Eisenberg & R. A. Fabes (Eds.). *Emotion and its regulation in early development: New directions for child development.* (55): 7–23. *The Jossey-Bass education series.* San Francisco: Jossey-Bass Publishers.
- Tsao, F-M., Liu, H-M., & Kuhl, P.K. (2004). Speech Perception in Infancy Predicts Language Development in the Second Year of Life: A Longitudinal Study. *Child Development*. (75)4: 1067–1084.

TRENDS IN EVOLUTION OF SIGNAL INTERPRETATION AS PRECURSORS OF THE ORIGIN OF HUMAN LANGUAGE

BURLAK SVETLANA^{*1}

*Corresponding Author: svetlana.burlak@bk.ru ¹Institute of Oriental Studies, RAS, Moscow, Russia

There are several major trends in signal interpretation in animal communication systems: from mere behaviour to specialized signals, from emotional signals to referential ones, from innate signals to arbitrary symbols; all of them are essential for the emergence of human language because human language is specialized and based on referential signs of symbolic nature. A need for a constant increase of the amount of symbols gave rise to phonology, syntax and other language-specific features. Focusing on signal interpretation makes it possible to construct a consistent model of language origin. It also provides a solution to the problem of honest communication emergence as well as a solution to the problem of gradual change.

1. Studies of animal communication systems

The communicative system of our non-human ancestors was an animal communication system, thus subject to the regularities that determine the evolution of such systems. Nevertheless, most language evolution models pay little attention to animal communication, and almost completely neglect trends that can be observed in its evolution.

Most scholars who study animal communication aim to identify the prerequisites of different aspects of human language capacity, as well as analogies between some traits in animal communication and human language. Thus, the ability to count is regarded as a prerequisite for recursion, preverbal concepts provide a basis for the development of language signs, birdsong syntax is considered as analogous to human language syntax (Okanoya 2002; Hurford 2012), geladas' lip-smacking is considered a precursor to speech (Bergman, 2013), and so on. In many works, animal communication systems are compared to human language in order to determine the distinctive features between the two (Hockett, 1960; Pinker & Jackendoff, 2008). For instance, N. Chomsky claims that the syntax (or, more specifically, the faculty of recursion) must have

played a cornerstone role in the language origin (Chomsky, 2002; Hauser *et al.*, 2014), and shows that non-human animals are incapable of acquiring a recursion-based grammar (Hauser *et al.*, 2002). T. Deacon (1998) claims that *Homo sapiens* is the only "symbolic species" (though he assumes that trained chimpanzees are capable of symbolization).

One author who pays more attention to animal communication systems is D. Bickerton (2009). According to his point of view, the human language emerged as a guidance system, like those that exist in several eusocial hymenopteran species (honey bees and some species of ants (Reznikova, 2017)), and its main goal was to convey information about remote objects. Nevertheless, even though D. Bickerton understands the importance of the evolutionary approach to the understanding of the emergence of human language, he says nothing about the evolutionary tendencies in animal communication. This creates a paradoxical situation: while the evolution of human body, its parts and organs is studied in details (see, e.g. Shubin (2008)), our communication system would appear to have sprung up out of nowhere, as if the body of our ancestors evolved and their communication system did not. This is, certainly, not the case. Communication systems of our ancestors did evolve as well, and there are several trends in their evolution that must have been essential for the emergence of human language.

K. Gibson (2010) pays attention to convergent evolution, because similar traits can evolve in very distantly related species in response to similar environments. She suggests that language-evolution process was jump-started by tool-assisted omnivorous extractive foraging.

Some authors present computational models of evolution of communication. For instance, T. Scott-Phillips (2010) shows that only that communicative system is evolutionary stable that is geared to maximize payoffs of both signaler and receiver. In his paper, however, nothing is said about evolutionary trends — directional changes of animal communication systems of one type to animal communication systems of another type.

Only a few authors, like I. Pepperberg (2012), mention a trend in animal communication system evolution in this respect (namely, the trend from innate to learned birdsong acquisition).

In this paper, we try to outline several patterns in the evolution of animal communication systems that may have played a role in human language evolution.

2. From Mere Behaviour to Signal

It is important to note that the properties of a communication system depend greatly on the way of life of the species (cf. Reznikova, 2017): only the most necessary things for fitness are encoded in signals. That way, signalling is an advantageous strategy in social species, but disadvantageous in solitary ones.

One of the features that oppose most animal communication systems to human language is that signals produced by animals are unintentional, and are not "designed from the beginning to communicate with. Rather they are modifications or stylizations or amplifications of things animals would do anyway" (Bickerton, 2009: 17). W. T. Fitch notes that, in animal communication, there are two levels of intentionality: "At the most basic level, so-called zero-order intentionality, the signaller has no intention psychologically, but the signal has nonetheless evolved to convey a message", while at the "first-order intentionality, a link between a mental representation and reality justifies our psychological interpretation of the signal, but implies no specific intent to inform another" (Fitch, 2010: 190).

It is important to emphasize that consciousness and intentionality are not necessary for the emergence of communication, only detectors are needed. An individual able to detect features of environment is eo ipso able to detect certain elements of conspecifics' appearance and behaviour; so when it is useful for some intentions (or observations) of an individual to be detected by conspecifics, the natural selection favours them to become more easily detectable. A good example is provided by two closely related species of toadhead agamas, reticulated toadhead agama Phrynocephalus reticulatus and sunwatcher toadhead agama *Phrynocephalus helioscopus* (Rogovin, 1991). For both species, it is important that a male does not spend its sperm on an already fecundated female. A female Ph. reticulatus may either bite a male in such case or run away. Conversely, for a female Ph. helioscopus, it is a bad strategy, because *Ph. helioscopus* have stronger jaws and are more purposeful, so a bite would be more dangerous for a male, and an attempt to run away is likely to fail. So, a female Ph. helioscopus produces a kind of press-up pantomime moving its body up and down, and a male ceases to pursue the female upon seeing this pantomime.

The "etymology" of this signal is quite transparent: it comes from common movements of two intentions, to bite and to run away, each of which can be seen in *Ph. reticulatus* behaviour. If a female *Ph. reticulatus* feels hesitation and spends a second or two choosing the behavioural program, we can see even both movements. However, there is a very important difference: in *Ph. reticulatus*,

they are merely movements, fluent and blurred, while in *Ph. helioscopus*, they function as a signal, they are longer, more distinct and more noticeable, being optimized for identification by male's detectors. Toadhead agamas' brain is very small, so signal movements cannot be blurred and fuzzy.

In ancestor toadhead agamas, natural selection favoured either the set "strong jaws + strength of purpose + signal" or the set "weaker jaws + lack of purposefulness + no signal". The males that failed to guess that a female with strong jaws had been already fecundated were either injured more seriously, or spent their sperm in vain in case they overtook the female and thus had less opportunity to produce offspring in both situations. The males of the species that is not so purposeful and does not have strong jaws can manage without guessing females' intentions since they would not be seriously injured in case a female chose to bite and most probably would not overtake a female that chose to run away.

Therefore, the main driving force for signals to develop is not the need to communicate something to somebody but the need to <u>recognize</u> something about a conspecific in order to minimize harm on oneself. For animals that are obliged to live close by their conspecifics, it is very important. It does not matter whether an individual producing a signal is even aware that the signal is produced, — if such signals help animals to minimize harm their conspecifics may inflict on them, they will be favoured by natural selection.

This is the solution of the problem of honest communication emergence (cf. Maynard Smith & Harper 2003): if individuals are selected not to transmit information but rather to understand it, than those that would notice only honest signals would never be deceived, would be able to choose a more appropriate behavioural program and produce more offspring as a consequence. This is also a solution for the problem of gradual change: what use is half a language? When we assume that the natural selection favours not production but interpretation, then any minor feature of behaviour that may give a cue to the perceiving individual would be useful because it would help the individual to obtain some information and thus to choose a more appropriate behavioural program.

For individuals whose state, mood and intentions are easily recognized by conspecifics, it is also of some use: e.g., if a conspecific notices the individual's aggressive mood, it may help avoid a conflict and avoid injures; if conspecifics of an individual are able to get information about food sources from it, such individual will have more satiated and healthy mates to transmit its genes with; an individual that will usually alarm its conspecifics in case of danger will have wider choice of potential mates. It is very important to have a wide choice because the genes of a potential mate need to be not "the best", but the most compatible with the genes of mating individual (Markov & Kulikov, 2006; Promislow *et al.*, 1998).

That way, it is not a mere chance that most signals produced by animals are not conscious and not deliberate: natural selection does not favour the production of signals, but their interpretation. The most ancient trend in animal communication evolution is emancipating certain elements of appearance and behaviour from needs of everyday routine and tuning them to be optimally recognized by brain detectors. Sound symbols of the human language may have also evolved in such a way, without any purpose or intention to make new signals.

3. From Emotional Signals to Referential Ones

The second important trend in evolution of animal communication is the evolution of referential signals from emotional ones. Emotional signals express the emotional state of an individual and are in factual connection to their objects. However, there are situations in which it is more useful to know not the emotion itself but the cause of it. The best example here is provided by alarm calls in vervet monkeys. Each call type elicits a different response (Seyfarth & Cheney 2012), and if a monkey that had heard a leopard alarm stayed on the ground looking at the sky (as if it had heard an eagle alarm), it would be eaten up; the same will happen to a monkey that mistook an eagle alarm for a leopard one. Therefore, natural selection favours the most precise discrimination between different types of alarm calls and production of the most distinct calls. There are several species having referential signals, namely: vervet monkeys (Cheney & Seyfarth, 1990; Seyfarth & Cheney, 2012; Hauser, 1996: 645-646), ground squirrels (Leger et al., 1980), marmots (Blumstein, 2007), meerkats (Manser & Bell, 2004; Hollén & Manser, 2006), ring-tailed lemurs (Macedonia, 1990; Pereira & Macedonia, 1991), Gunnison's prairie dogs (Slobodchikoff et al., 1991) and even chicken (Evans & Evans, 1999). There is a hypothesis according to which referential signals may have a very precise meaning, for example, denoting not only the type of object but also its colour (Slobodchikoff et al., 2009).

There is a crucial difference between emotional and referential signals. Emotional signals may vary along a continuum because the degree of feeling is continuous, while referential signals are discrete because there is no intermediary between, say, an eagle and a leopard. Humans use signals of both types: our morphemes, words and sentences are referential while elements of nonverbal communication are emotional signals.

In some works, referential signals are considered an intermediary stage on the way from an animal communication system to human language: at first, there was a few referential signals, as in vervet monkeys, then they grew in number (Bickerton, 2003: 79). However, this is certainly not the case for two reasons. First, human language is an open system, we are capable of producing an infinite number of sentences and even of words. On the contrary, vervet monkeys' referential signals are innate, and such a system can never become an infinite one. Second, apes, which are the closest relatives of humans, do not have referential alarm calls. Thus, in human clade, the ability to use referential signals must have been formed *de novo* (and that is why it is not very similar to what we see in vervet monkeys).

Evolution from emotional signals to referential ones can be traced in ground squirrels. Evolutionary primitive ground squirrels have sounds reflecting the degree of their fear: when fear is great, they produce whistles, when fear is weak they produce trills, and in intermediary cases intermediary sounds break from them. In evolutionary more advanced species, alarm calls become specialized: whistles break from the individuals noticing a flying raptor and chatter-chat calls (homologous to trill) are elicited by terrestrial predators, and no intermediary variant exists. The correlation between the specialization of a species and the number of specialized signals it uses can be seen in martmots (see Blumstein, 2007: 373).

Sometimes it is not easy to determine whether a signal is emotional or referential. Food calls in chimpanzees (Slocombe & Zuberbühler, 2005) may refer to outer world objects, but they may be emotional signals as well reflecting not the outer world objects themselves but merely the emotions of vocalizing individuals. However, when choice is limited emotional signals may suffice to choose the most appropriate behavioural program. So, even if great apes completely lack referential signals that could evolve into words of human language, it is not an obstacle, because referential signals can evolve (and, in fact, they do evolve) from emotional ones. Therefore, even if the earliest signals in hominines were emotional, a well-known evolutionary mechanism would turn them into referential ones.

4. Decreasing of the Rate of Innateness

The most important trend in animal communication systems' evolution for the origin of language is a decrease in the rate of innateness in signals. In the most primitive communication systems, both signifier and signified are innate. A good example is provided by a male three-spined stickleback. A male stickleback is not aware that its belly becomes red and is unable either to change its colour or to suppress aggressive reaction. Sticklebacks are not very clever, and this mechanism helps them to mate their females and not to let other males mate females instead of them. Natural selection favours those males that attack red-bellied sticklebacks and court silver-bellied ones during the mating period, irrespective of any awareness.

The next stage consists in so-called "hierarchical signals" that have innate form and learned meaning. For example (Fridman, 2013: 8), great spotted woodpeckers (*Dendrocopos major*) when they defend their territories show their outermost rudders (tail feathers) to their neighbours. Pattern of white spots on the tail feathers is unique for each individual, and the birds recognize their neighbours by sight and remember the results of previous territorial conflicts. That way, when a woodpecker recognizes the neighbouring individual it becomes able to choose its behavioural program based on the information about hierarchical relations between them. In this case, signifier (the pattern of spots) is innate while the signified (information about hierarchical relations) is not.

The next stage consists in so-called *ad-hoc*-signals (Barulin, 2002). Species having most advanced cognitive abilities, e.g. chimpanzees (Goodall, 1986), can interpret behaviour of conspecifics and are therefore able to turn its elements into signals when it is necessary. In different groups of chimpanzees, different signals are used, which proves cultural transmission (Whiten *et al.*, 1999).

This trend shows how the communication becomes increasingly connected with individual experience and cognition. The human language is the next step in this row: once invented, *ad-hoc* signals are learned and transmitted culturally.

5. Human Language as a Communication System Suited for Commenting

An important tendency in evolution of the communication systems in Primates was revealed by M. A. Deryagina and S. V. Vasiliev (1993): the most ancestral signals denote aggressive intentions, and in more advanced taxa the number and percentage of friendly signals increases (and friendly communicative complexes are formed). It is very important because cooperation plays crucial role in the origin of human language (cf. Tomasello, 2008).

Those aspects of language capacity that are uniquely human — combinatorial phonology, derivational and inflectional morphology, hierarchical syntactic structures, compositional semantics, discourse-organizing means,

speech registers and styles, connotations and so on (see Pinker & Jackendoff, 2008) — can be carried out only in a communicative system containing a vast (potentially infinite) number of signals (Burlak, 2011: 82). For a communicative system having, say, ten signals, a set of phonological oppositions or a hierarchical syntactic structure is unnecessary.

Therefore, the main line of language emergence was the accumulation of signs, and the main challenge of human evolution was to learn more and more communicative signals. Where does such a challenge come from? The ancestors of ancient hominids were primates with a considerably high intelligence level, able (and willing) to guess reasons, draw conclusions and adopt behavioural patterns (with the help of mirror neurons, see (Arbib, 2012)). During the course of human evolution, these abilities developed — endocasts show the growth of prefrontal cortex in human lineage (Drobyshevskij, 2007). As noted by S. Pinker, language emergence was a part of a complex adaptation to the cognitive niche (Pinker, 2003).

Ecological niche of a savannah big terrestrial omnivorous animal demanded a constant augmentation of number of behavioural patterns and number of observable environmental details useful for guessing reasons and drawing conclusions. For gregarious primates, there was a need to communicate all this to conspecifics. Those who managed to share their findings with their relatives were more efficient in bringing up their offspring. Thus, natural selection favoured the hominid groups whose members were better at sharing knowledge with each other. When hominids learned how to make tools, the number of their behavioural patterns increased, and so did the number of objects and details worth noticing. That way, a means was needed to sort them out, to figure out, which behaviour is best suited to in particular circumstances. Natural selection thus favoured the groups in which a fact known to one individual became known to others easily. When an individual noticed something useful for choosing an optimal behavioural pattern in a situation, and drew his/her relative's attention to this, both would benefit from it. An individual who reported what (s)he saw or heard to his/her relatives became an additional pair of eyes and ears for each member of the group. Note, that such reports did not need to be either conscious or articulate, they only had to be interpreted by others and to provide them with material allowing to draw appropriate conclusions. Unintentional and inarticulate "signals" may be interpreted as well. The main driving force of the evolution of hominids' communicative system was thus a great need for such reports (Burlak, 2012a).

People — even now — are instinctively convinced that language is a system for naming: "words" are, first of all, names of things, while "sentences" are, first of all, comments to situations. If a non-linguist is asked to say some "words" in another language, (s)he will most probably come up with some names of objects (not verbs, or adjectives, or particles, or even abstract nouns). When one wants to deal with a "sentence", (s)he will at first consider narrative sentences: *The farmer killed the duckling, Colorless green ideas sleep furiously* and so on. "Basic word order" of a language is nothing more than the word order normally used in narrative sentences. Naturally, such sentences are far from being the only type of utterances occurring in our everyday speech.

Language is not a thing in itself, psychological experiments show that when people are asked to learn certain sentences, they tend to store in their long-term memory not the sentences they were asked to learn but the conclusions that are based on them (Anderson, 2015: 161-162).

At first, such comments, probably, were like modern children's private (egocentric) speech (Vygotsky, 1986). When a child plays, (s)he is constantly telling him/herself what (s)he sees or hears, does, needs, or is going to do. Although such speech is called private (egocentric), children seem to address it not only to themselves, but also to others: in presence of deaf people or foreigners they tend to play almost silently (ibid.). When a child grows up, egocentrical speech turns into inner speech. Human language, as a whole, probably, also evolved in such a way.

6. Emergence of an Open Communication System

When a strong demand for comments arose, the natural selection began to favour the groups consisting of individuals who were effective not only in interpreting comments of conspecifics but also in producing comments clear enough to be understood properly. For these purposes, it is optimal to know more signals. So natural selection favoured the groups in which the offspring was willing to know names of all the things in the world.

First signals were certainly iconic, because they had to be understood without any previous knowledge, but after some repetitions, when a signal became known to several members of a group, it could lose its iconicity (as a result of habituation) and become a symbol. In such symbols, distinctive features were singled out, which could be used to distinguish one symbol from another. This rendered it possible to learn new signals very quickly. When multiple signals exist, associations emerge between them, because a word having a meaning activates a neuron assembly, and some of its neurons also participate in coalitions corresponding to other meanings; these meanings become associated with each other. This provided the communicative system with an additional function: the one to tell about objects and situations that are beyond the speaker's immediate experience.

When many signals are known, it became possible to invent new signs not on the base of a specific real situation but on the base of previously known signs by means of modifying them. Signs can be modified in any possible way: by adding vocalization to a gesture, or by adding new sounds to a vocalization, by changing the intonation, by performing a given gesture with the other hand, etc. When several pairs of signs and their modified variants became known to the members of a hominid group, they gained the possibility to generalize modifications. That was, obviously, a crucial point of the language evolution because it made the entire communicative system constructable: an individual who knew a few signals and several modification rules became able to produce a great number of new understandable utterances.

The need for comments was also a driving force of switching from gestures to a sound communication system. When the main goal of communication is to make a certain conspecific do something, a communication system can be based on gestures: an individual would approach a conspecific and begin to communicate. But when a communication system is designed to share attention with conspecifics (Tomasello, 2008), it will inevitably become acoustic: an individual ceasing his/her activity to begin a communication would stop perceiving the things (s)he needs to communicate.

For an acoustic communication system with an infinite number of potential utterances, a complex of features is needed. The most relevant features for articulate speech are lowered larynx and increased spinal canal width (indicating the increased capacity of breath control). Another feature of great importance is the lack of air sacs (Boer, 2011). When articulated speech arose, hearing abilities must have increased (chimpanzees have a lower ability to hear high frequency sounds (see Martínez *et al.*, 2008)). The full complex of adaptations for articulated speech appears first in *Homo heidelbergensis*: this species lacked air sacs, had a wider spinal canal and bigger brain, and maintained a relatively high sensitivity from 2 kHz up to 4 kHz (an interval that contains acoustic features that can be modified by articulation), which renders it highly probable that *Homo heidelbergensis* was capable of articulated speech (Burlak, 2012b, Dediu & Levinson, 2013). However, there is no reason to believe that this species had a full-fledged human language, because many aspects of the latter

are closely associated with frontal lobes, which are smaller in *Homo* heidelbergensis than in *Homo* sapiens.

Thus, the increasing demand for incoming information formed a communication system with a potentially endless number of possible utterances. In turn, this created a demand for phonology, syntax and other language-specific features. Focusing on signal interpretation makes it possible not only to solve the problem of honest communication emergence and the problem of gradual change but also to construct a consistent model of language origin.

References

- Anderson, J. R. (2015). *Cognitive psychology and its implications* (8th ed.). New York: Worth Publishers.
- Arbib, M. A. (2012) Mirror systems: Evolving imitation and the bridge from praxis to language. In M. Tallerman and K. R. Gibson (Eds.) *The Oxford Handbook of Language Evolution* (pp. 207-215). Oxford: Oxford Univ. Press.
- Barulin, A. N. (2002). Osnovanija semiotiki: znaki, znakovyje sistemy, kommunikatsija. Part 1. Moscow: Sport i kul'tura – 2000. (Foundations of the semiotics: Signs, sign systems, communication.)
- Bickerton, D. 2003. Symbol and structure: A comprehensive framework for language evolution. In M. H. Christiansen and S. Kirby (Eds.), *Language evolution* (pp. 77-93). Oxford: Oxford University Press.
- Bickerton, D. (2009). Adam's tongue: how humans made language, how language made humans. New York: Farrar, Straus and Giroux.
- Bergman, T. J. (2013). Speech-like vocalized lip-smacking in geladas *Current biology*, 23(7), R268–R269.
- Blumstein, D. T. (2007) The evolution, function, and meaning of marmot alarm communication. *Advances in the Study of Behavior*, *37*, 371–400.
- Boer, B. de (2011). Loss of air sacs improved hominin speech abilities. *Journal* of Human Evolution, 62(1), 1–6.
- Burlak, S. A. (2011). *Proiskhozhdenije jazyka: Fakty, issledovanija, gipotezy.* Moscow: CORPUS. (The origin of language: Facts, research, hypotheses.)
- Burlak, S. A. (2012a). Chelovecheskij jazyk kak kommunikativnaja sistemakommentarij Vestnik Rossijskogo Gosudarstvennogo gumanitarnogo universiteta, 88, 132–145. (Human language as a communicative system for commenting.)
- Burlak, S. A. (2012b). Vremya pojavlenija zvuchashchej rechi po dannym antropologii. Vestnik Moskovskogo gosudarstvennogo universiteta, Ser. 23 Antropologija, 3, 109–119. (When did articulate speech arise (according to anthropological data)?)

- Cheney, D. L., & Seyfarth, R. M. (1990). *How monkeys see the world*. Chicago and London: University of Chicago Press.
- Chomsky, N. (2002). *On nature and language*. Cambridge, England: Cambridge University Press.
- Deacon, T. (1998). *The symbolic species: The co-evolution of language and the human brain*. London: Penguin Books.
- Dediu, D., & Levinson, S. C. (2013). On the antiquity of language: The reinterpretation of Neandertal linguistic capacities and its consequences. *Frontiers in Psychology*, 4, 397.
- Deryagina, M. A., & Vasil'jev, S. V. (1993). Formy obshchenija u primatov i proiskhozhdenije chelovecheskogo jazyka. In O. A. Donskikh (Fd.) Jazyk v okeane jazykov (pp. 60–85). Novosibirsk: Sibirskij khronograf. (Communication forms in Primates and human language origin.)
- Drobyshevskij, S. V. (2007). Evolutsija mozga cheloveka: Analiz Endocraniometricheskikh Priznakov gominid. Moscow: URSS. (The evolution of human brain: an analysis of endocast-measurment features of hominines.)
- Evans, Ch. S., & Evans, L. (1999). Chicken food calls are functionally referential. *Animal behaviour*, 58(2), 307–319.
- Fitch, W. T. (2010). *The evolution of language*. Cambridge: Cambridge Univ. Press.
- Fridman, V. S. (2013). Ot stimula k simvolu: Signaly v kommunikatsii pozvonochnykh. Part 2: Znak v kommunikatsii zhivotnykh. Evolutsija sistem signalizatsii pozvonochnykh. Metodologija analiza signal'nykh sistem. Moscow: Knizhnyj dom "LIBROKOM". (From stimulus to symbol: signals in vertebrates' communication. Part 2: Sign in animal communication. Evolution of signaling systems in vertebrates. Methodology of signaling systems' analysis.)
- Gibson, K. R. (2010). Talking about birds, bees, and primates, too. Implications for language evolution. In A. D. M. Smith, M. Schouwstra, B. de Boer, & K. Smith (Eds.) The Evolution of Language. Proceedings of the 8th International Conference (EVOLANG8), Utrecht (pp. 153–159). London: World Scientific.
- Goodall, J. (1986). *The chimpanzees of Gombe: Patterns of behavior*. Cambridge, MA: The Belknap Press of Harvard University Press.
- Hauser, M. D. (1996). *The evolution of communication*. Cambridge, MA: MIT Press.
- Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The Faculty of language: What is it, who has it, and how did it evolve? *Science*, 298, 1569–1579.
- Hauser, M. D., Yang, C., Berwick, R. C., Tattersall, I., Ryan, M. J., Watumull, J., Chomsky, N., & Lewontin, R. C. (2014). The mystery of language evolution. *Frontiers in psychology*, 5(401), 1–12.
- Hockett, Ch. F. (1960). The origin of speech. Scientific American, 203, 89-97.

- Hollén, L. I., & Manser, M. B. (2006). Ontogeny of alarm call responses in meerkats, *Suricata suricatta*: the roles of age, sex and nearby conspecifics. *Animal behaviour*, 72 (6), 1345–1353.
- Hurford, J.R. (2012). *The Origins of Grammar: Language in the Light of Evolution*. Oxford: Oxford University Press.
- Leger, D. W., Owings, D. H., & Gelfand, D. L. (1980). Single-Note Vocalizations of California ground squirrels: Graded signals and situationspecificity of predator and socially evoked calls. *Zeitschrift für Tierpsychologie*, 52, 227–246.
- Macedonia, J. M. (1990). What is communicated in the antipredator calls of lemurs: Evidence from playback experiments with ringtailed and ruffed lemurs. *Ethology*, 86, 177–190.
- Manser, M. B., & Bell, M. B. (2004). Spatial representation of shelter locations in meerkats *Suricata suricatta*. *Animal behaviour*, 68, 151–157.
- Markov, A. V., & Kulikov, A. M. (2006). Gipoteza immunologicheskogo testirovanija partnerov — soglasovannosť razvitija adaptatsij i smeny polovykh predpochtenij *Izvestija RAN, Ser. biologicheskaja*, 3, 261–274. (A hypothesis about immunological testing of mating partners: coherence of the development of adaptations and change of sexual preferences.)
- Martínez, I., Quam, R., Rosa-Zurera, M., Arsuaga, J. L. (2008). Auditory capacities of human fossils: A new approach to the origin of speech. *The Journal of the Acoustical Society of America*, 123(5), 3606.
- Maynard Smith, J., & Harper, D. G. C. (2003). *Animal signals*. Oxford: Oxford University Press.
- Okanoya, K. (2002). Sexual display as a syntactical vehicle: The evolution of syntax in birdsong and human language through sexual selection. In A. Wray (Ed.), *Transition to language* (pp. 46–63). Oxford: Oxford University Press.
- Pepperberg, I. M. (2012). Evolution of Communication and Language: Insights From Parrots and Songbirds. In M. Tallerman and K. R. Gibson (Eds.) *The Oxford Handbook of Language Evolution* (pp. 109–119). Oxford: Oxford Univ. Press.
- Pereira, M. E., & Macedonia, J. M. (1991). Ringtailed lemur antipredator calls denote predator class, not response urgency *Animal Behaviour*, 41, 543-544.
- Pinker, S. 2003. Language as an adaptation to the cognitive niche. In M. H. Christiansen and S. Kirby (Eds.), *Language evolution* (pp. 16–37). Oxford: Oxford University Press.
- Pinker, S., & Jackendoff, R. (2008). The components of language: What's specific to language, and what's specific to humans. In M. H. Christiansen, C. Collins and S. Edelman (Eds), *Language Universals* (pp. 126–152). Oxford: Oxford University Press.
- Promislow, D. E. L., Smith, E. A., & Pearse, L. (1998). Adult fitness consequences of sexual selection in *Drosophila melanogaster*. *Proceedings* of the National Academy of Science, USA, 95(18), 10687–10692.

- Reznikova, Zh. I. (2017). Studying Animal Languages Without Translation: An Insight from Ants. Cham: Springer.
- Rogovin, K. A. (1991). Sotsial'noje povedenije kruglogolovok *Phrynocephalus helioscopus* i *Ph. reticulatus* (Reptilia, *Agamidae*) i ikh vzaimootnoshenija v sovmestnykh poselenijakh. *Zoologicheskij zhurnal*, 70(3), pp. 61–72. (Social behaviour of toadhead agamas *Phrynocephalus helioscopus* and *Ph. reticulatus* (Reptilia, *Agamidae*) and their interrelations in common settlements.)
- Scott-Phillips T. C. The evolution of communication and relevance. In A. D. M. Smith, M. Schouwstra, B. de Boer, & K. Smith (Eds.) *The Evolution of Language. Proceedings of the 8th International Conference (EVOLANG8), Utrecht* (pp. 489–490). London: World Scientific.
- Seyfarth, R. M., & Cheney, D. L. (2012). Primate social cognition as a precursor to language. In M. Tallerman and K. R. Gibson (Eds.) *The Oxford Handbook* of Language Evolution (pp. 59-70). Oxford: Oxford Univ. Press.
- Shubin, N. (2008). Your Inner Fish: A Journey into the 3.5-Billion-Year History of the Human Body. New York: Pantheon Books.
- Slobodchikoff, C. N., Kiriazis, J., Fischer, Ch., & Creef, E. (1991). Semantic information distinguishing individual predators in the alarm calls of Gunnison's prairie dogs. *Animal Behaviour*, 42(5), 713–719.
- Slobodchikoff, C. N., Paseka, A., & Verdolin, J. L. (2009). Prairie dog alarm calls encode labels about predator colors. *Animal Cognition*, 12, 435–439.
- Slocombe, K. E.; Zuberbühler, K. (2005). Functionally Referential Communication in a Chimpanzee. *Current Biology*, *15*(*19*), 1779–1784.
- Tomasello, M. (2008). *Origin of human communication*. Cambridge (Mass.); London: The MIT Press.
- Vygotsky, L. S. (1986). *Thought and Language*. Cambridge (Mass): The MIT Press.
- Whiten, A., Goodall, J., McGrew, W. C., Nishida, T., Reynolds, V., Sugiyama, Y., Tutin, C. E. G., Wrangham, R. W., & Boesch, Ch. (1999). Cultures in chimpanzees. *Nature*, 399, 682–685.

INTERACTIVE SEQUENCES MODULATE THE SELECTION OF EXPRESSIVE FORMS IN CROSS-SIGNING

Kang-Suk Buyn*1, Connie de Vos2, Seán G. Roberts3, and Stephen C. Levinson1

*Corresponding Author: Kang-Suk.Byun@mpi.nl ¹Language and Cognition Department, Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands ²Center for Language Studies, Radboud University Nijmegen, Netherlands ³excd.lab. University of Bristol, UK

This study looks at how linguistic conventions arise in the context of face-toface, real-time interaction. This topic is difficult to study because initial contact events for most languages happened long ago. Two approaches try to tackle this. The first is experimental semiotics: putting people in a situation where they need to improvise a new communication system. Studies have considered the relative role of different biases in this process (Tamariz, et al., 2014) such as frequency bias (where the more frequent signs prevail), content bias (where the more iconically - motivated, and/or easily - articulated form is selected), and coordination bias (where participants attempt to match each other's usage). However, these experiments typically involve an artificial language or a restricted or unfamiliar communication medium which are used by participants who already share a common language and culture. This limits the ecological validity of these experiments, especially relating to face-to-face interaction.

Another approach has been to study the emergence of new signed languages which emerge spontaneously from scratch, allowing the study of the formation processes of modern human languages in real life. This process has been well-documented in the case of Nicaraguan Sign Language, which emerged over the course of several decades in a deaf school (Senghas & Coppola, 2001), as well as in multiple 'deaf villages' where a local sign language has emerged from the interaction of deaf and hearing community members (Meir et al. 2010). Lesser-known instances of *de novo* signed communication arises between deaf and hearing non-signers (trans-languaging), and among deaf signers who do not know a common written or signed language (cross-signing, Kusters et al. 2017; Buyn et al., 2017). Cross-signing is of particular interest as it creates a real-time pressure to establish a shared communicative repertoire. However, these studies are often not designed to be experimentally controlled, and rarely capture the very first period of the emergence of a signed language.

In this study we combine the control of experimental semiotics with the eco-

logical validity of cross-signing in order to assess the relative contribution of different biases to the emergence of a shared lexicon. We recreated the context for cross-signing to emerge in a lab by flying deaf individuals from Nepal, Jordan and Indonesia to India and documenting their first encounters. This resulted in a 320-minute video corpus between dyads of signers of Nepali Sign Language, Indian Sign Language, Jordanian Sign Language and Indonesian Sign Language (these languages are mutually unintelligible and none of the signers knew any of the other languages). Recordings took place at the first meeting, after one week of living together, and after three weeks and included both spontaneous conversations and structured communication tasks, including the spot-the-differences task reported here. In this task, two signers were given a cartoon image each which differed only in the way certain objects were coloured. They could not see their partner's image. Their task was to identify the differences in colour by communicating spontaneously face-to-face. We coded and examined the various colour expressions exploited by the participants. The pairs repeated the experiment with new images after 1 week and after 3 weeks.

Participants initially used a range of strategies, including pointing, articulating signs for common objects with that colour (e.g. referring to a common iconic sign for 'tree' and pointing to the base to mean 'brown'), and their own native variants. However, after three weeks a consensus has been formed. For example, every-one uses the Indian signer's variant for 'green' and the Nepali signer's improvised 'tree-trunk' variant for 'brown'. Sequential analyses indicate that signers continuously assess the relative ease with which their forms might be understood, and adopt interactional strategies (e.g. try-marking, repair sequences) to manage communicative difficulties that arise. Mixed effects modelling was used to analyse which factors promoted the selection of particular signs. There were significant effects for frequency, content and coordination biases, supporting the findings of previous experimental studies, but the interactional context in which a form was used also mattered. This ability to study the very early stages of language emergence in a real yet controlled environment is an exciting addition to language evolution research approaches.

Acknowledgements

SGR supported by a Leverhulme early career fellowship to SGR (ECF-2016-435) and an European Research Council Advanced Grant No. 269484 INTERACT to SCL.

References

Byun, K.-S., De Vos, C., Bradford, A., Zeshan, U., & Levinson, S. C. (2017). First encounters: Repair sequences in cross-signing. *Topics in Cognitive Science, doi:10.1111/tops.12303.*

- Kusters, A., Spotti, M., Swanwick, R., & Tapio, E. (2017). Beyond languages, beyond modalities: transforming the study of semiotic repertoires. *International Journal of Multilingualism*, 1–14.
- Meir, I., Sandler, W., Padden, C., Aronoff, M., et al.. (2010). Emerging sign languages. Oxford handbook of deaf studies, language, and education, 2, 267–280.
- Senghas, A., & Coppola, M. (2001). Children creating language: How nicaraguan sign language acquired a spatial grammar. *Psychological science*, 12(4), 323–328.
- Tamariz, M., Ellison, T. M., Barr, D. J., & Fay, N. (2014). Cultural selection drives the evolution of human communication systems. *Proceedings of the Royal Society of London B: Biological Sciences*, 281(1788), 20140488.

CONCEPTUAL STRUCTURE IS SHAPED BY COMPETING PRESSURES FOR SIMPLICITY AND INFORMATIVENESS

Jon W. Carr^{*1}, Kenny Smith¹, Jennifer Culbertson¹, and Simon Kirby¹

*Corresponding Author: j.w.carr@ed.ac.uk ¹Centre for Language Evolution, University of Edinburgh, UK

Languages are shaped by competing pressures from learning and communication. Learning favours simple languages, while communication favours informative ones, giving rise to the simplicity–informativeness tradeoff. Languages that evolve under this tradeoff are both maximally simple (learnable) and maximally informative (communicatively useful). This has been shown in natural language and in experimental settings. For example, Kemp and Regier (2012) showed that kinship systems exist at the optimal frontier of simplicity and informativeness. In a separate line of experimental work, Kirby, Tamariz, Cornish, and Smith (2015) showed that when artificial languages evolve under a learning pressure alone, they become simple and uninformative, and when languages evolve during communication, they become complex and informative; it is only when both pressures are at play that we find languages at the optimal frontier.

However, a recent iterated learning experiment by Carstensen, Xu, Smith, and Regier (2015) showed that artificial languages expressing spatial relationships tended to become more informative when subjected to a pressure from learning. This is a surprising result given the previous work briefly reviewed above, which says that informativeness is driven by the pressure from communication, not from learning. One potential explanation for this result lies in their measure of informativeness, *communicative cost*, which is sensitive to (a) the number of words that the language is comprised of (expressivity) and (b) the extent to which similar meanings are expressed by the same word (which we will term convexity). In their experiment expressivity was fixed at four words. As a result, the reduction in communicative cost they found must be due to categories evolving to become more convex, i.e. picking out increasingly tightly-clustered sets of meanings.

To demonstrate that learning favours convex categories, we conducted two experiments in which participants learned and produced a category system for stimuli varying on two dimensions, size and angle. In Experiment 1 participants were trained on one of three systems: One marking a distinction in angle, one marking a distinction in size, and one marking a distinction on both dimensions (see Fig. 1). The results indicated that the Angle-only system was easiest to learn,

followed by the Size-only system; the Angle & Size system was hardest to learn, despite having the lowest communicative cost.



Figure 1. The three systems participants were trained on in Experiment 1. Each colour indicates a category (a group of meanings labelled by the same word).

In Experiment 2, the output of one participant became the input to the following participant in a standard iterated learning design. An example chain is shown in Fig. 2. Over 12 generations the category system became increasingly easy to learn, as indicated by decreasing intergenerational transmission error. Furthermore, in the majority of chains, the language converged on a system marking only a distinction on the angle dimension, which participants found easiest in Experiment 1. This increase in simplicity is driven by increasing convexity.



Figure 2. An example iterated learning chain from Experiment 2, showing 12 generations of change in the category system. The initial state of the language at generation 0 was generated randomly. By generation 12, the language has evolved to mark an angle distinction.

We also found that most chains converged on fewer than four categories. This suggests that iterated learning acts as a pressure for simplicity by simultaneously decreasing expressivity and increasing convexity. However, if, as in Carstensen et al. (2015), expressivity is held constant, the learning pressure can only act through convexity: Although languages may become more informative under iterated learning, they do so not because of a pressure to be more communicatively useful, which in Carstensen et al.'s study necessarily decreases communicative cost as a side-effect of increasing convexity. This therefore suggests that, contra Carstensen et al. (2015), languages which are both simple yet informative will only emerge when pressures from learning and communication are at play. We support these conclusions with a Bayesian iterated learning model that displays strikingly similar results.

Acknowledgements

JWC was funded by the Economic and Social Research Council (grant number ES/J500136/1).

References

- Carstensen, A., Xu, J., Smith, C. T., & Regier, T. (2015). Language evolution in the lab tends toward informative communication. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.), *Proceedings of the 37th Annual Conference of the Cognitive Science Society* (pp. 303–308). Austin, TX: Cognitive Science Society.
- Kemp, C., & Regier, T. (2012). Kinship categories across languages reflect general communicative principles. *Science*, 336, 1049–1054.
- Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, 141, 87–102.

ICONICITY AND CONVENTION IN THE MANUAL MODALITY: PANTOMIME IN LANGUAGE ORIGINS

ERICA A. CARTMILL

cartmill@anthro.ucla.edu Department of Anthropology, University of California, Los Angeles, USA

Iconicity features prominently in manual communication systems, from cospeech gesture to fully-structured sign languages. Despite early reports to the contrary (Orlansky & Bonvillian, 1984), there is growing evidence that iconicity aides in the early acquisition of signed languages (Thompson et al., 2012), though older children benefit more (Magid & Pyers, 2017) and there may be different effects in second language learning (Ortega, 2017). Of course, iconicity is not limited to the gestural realm. Recent studies show that spoken languages contain many iconic elements, particularly in child speech (Monaghan et al., 2014; Perry et al., 2017). In sum: iconicity is prevalent in language, and may even be a driver of linguistic form (Dingemanse et al., 2015).

Iconicity is important to discussions of the evolution of language because iconic communication could, in principle, be understood without a shared linguistic system (i.e., without language). "Gesture first" theories of language origins emphasize the importance of iconicity as a bridge to language, arguing that iconic gesture could have provided an advantage to human ancestors before the emergence of language (e.g., Zlatev, 2008; Arbib, 2015). These authors propose that iconic gesture in the form of *pantomime* formed a transitional stage in the evolution of symbolic communication, between the ability to imitate actions and the emergence of communicative conventions. This is a compelling story, but I argue that iconicity (especially pantomime) did not play a substantial role in language origins until well after the emergence of conventional systems.

Despite the prevalence of iconic words and signs in modern languages, iconicity of the sort required by these theories is not a trivial matter understanding and producing iconic gestures requires complex mental representation and sophisticated analogic reasoning. This may put it beyond reach of very young children and non-human primates, which raises questions

about whether iconic gesture and pantomime could be stepping-stones to symbolic, codified language.

Recent work on iconicity has focused on identifying iconic features in conventional languages (like spoken English or ASL). However, it might reveal more about the underlying cognitive demands to study non-conventionalized systems (e.g., co-speech gesture or homesign). These systems do not primarily rely on shared conventions and may better reflect children's own understanding of iconicity. Hearing children do not begin using iconic gestures until relatively late in development, after they have mastered both conventional and indexical gestures. Use of iconic gesture undergoes a rapid period of growth around 2.5 years of age (Özçalişkan & Goldin-Meadow, 2011). However, little is known about the cause of this growth, particularly whether it is driven by cognitive or communicative constraints. If understanding iconicity is indeed a complex task, then we might expect it to be more closely linked to cognitive development.

Using data from a recent study comparing the development of manual iconicity in co-speech gesture and homesign over the first 5 years of life (Cartmill et al., 2017), I explore the relevance of iconic gesture development for theories proposing pantomime as an evolutionary stepping-stone to conventional language. Homesign is vastly different from co-speech gesture, and has many structural properties of language (unlike gesture). Nevertheless, homesigners and hearing children show many similarities in the ways they begin to use their hands to reference the world iconically without conventional models of appropriate sign forms, particularly in the way the hand maps to aspects of objects or events (e.g., is the gesturing hand an hand or an object). These similarities suggest that the ability to use iconicity in the manual modality may develop in similar ways in both gesturing and homesigning children, supporting the theory that the growth spurt in iconic gesture has more to do with cognitive development than with the particular features of the linguistic system.

This similarity has direct implications for gesture first theories. While iconicity may help children learn words in conventional (spoken or signed) languages, the linguistic framework may provide crucial support for correctly interpreting the iconic features of unknown words. Manual iconicity without the support of conventional manual structures may not confer the same advantages. The presence of sound-symbolic words and iconic features of signed languages may thus tell us little about the likelihood of a pantomimic protolanguage.

References

- Arbib, M. A. (2012). *How the brain got language: The mirror system hypothesis* (Vol. 16). OUP USA.
- Cartmill, E. A., Rissman, L., Novack, M. A., & Goldin-Meadow, S. (2017). The development of iconicity in children's co-speech gesture and homesign. *Language, Interaction and Acquisition*, 8(1), 42-68.
- Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. *Trends in Cognitive Sciences*, 19(10), 603-615.
- Magid, R. W., & Pyers, J. E. (2017). "I use it when I see it": The role of development and experience in Deaf and hearing children's understanding of iconic gesture. *Cognition*, 162, 73-86.
- Monaghan, P., Shillcock, R. C., Christiansen, M. H., & Kirby, S. (2014). How arbitrary is language? *Phil. Trans. R. Soc. B*, 369(1651), 20130299.
- Orlansky, M. D., & Bonvillian, J. D. (1984). The role of iconicity in early sign language acquisition. *The Journal of Speech and Hearing Disorders*, 49(3), 287-292.
- Ortega, G. (2017). Iconicity and Sign Lexical Acquisition: A Review. *Frontiers in Psychology*, 8.
- Özçalişkan, Ş., & Goldin-Meadow, S. (2011). Is there an iconic gesture spurt at 26 months? *Integrating gestures: The interdisciplinary nature of gesture. Amsterdam, NL: John Benjamins.*
- Perry, L. K., Perlman, M., Winter, B., Massaro, D. W., & Lupyan, G. (2017). Iconicity in the speech of children and adults. *Developmental Science*.
- Thompson, R. L., Vinson, D. P., Woll, B., & Vigliocco, G. (2012). The road to language learning is iconic: Evidence from British Sign Language. *Psychological science*, 23(12), 1443-1448.
- Zlatev, J. (2008). The co-evolution of intersubjectivity and bodily mimesis. *The shared mind: Perspectives on intersubjectivity*, 215-244.

BEAT PERCEPTION IN A NON-VOCAL LEARNER: RATS CAN IDENTIFY ISOCHRONOUS BEATS

Alexandre Celma-Miralles^{*1}, and Juan M. Toro^{1,2}

*Corresponding Author: alexandre.celma@upf.edu
 ¹ Center for Brain and Cognition - ETIC, Universitat Pompeu Fabra, Barcelona, Spain
 ² Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain

The origins of rhythm in music and language seem to be intertwined and may share evolutionary pathways (Fitch, 2006, 2013; Patel, 2006, 2010; Ravignani and Madison, 2017). The development of song in birds and speech in humans follows similar sensorimotor phases and parallel periods of vocal learning (Marler, 1970, 1976), an ability that allows animals to imitate and modify the vocalizations learned from other individuals thanks to specific neural connections (Jarvis, 2006, 2007). For the last decade, the vocal learning beat perception and synchronization hypothesis (Patel, 2006; Patel, Iversen, Bregman, and Schulz, 2009; Schachner, Brady, Pepperberg and Hauser, 2009) has been widely accepted among researchers, proposing that only vocal learners can extract and entrain to a periodic pulse. However, current animal studies reported beat entrainment (Cook, Rouse, Wilson and Reichmuth, 2013) and auditory synchronization (Hattori, Tomonaga and Matsuzawa, 2013) in not classically classified vocal learners, and non-human primates display some rhythmic behaviors in social interactions (see Merchant and Honing, 2013; Ravignani, Gingras, Asano, Sonnweber, Matellán and Fitch, 2013).

To explore to what extent beat perception appears in non-vocal learners, we tested rats (Rattus norvegicus) in a go-no go paradigm. Several studies on comparative research have successfully used this paradigm with rats to discern what components of music and language are actually shared with other species (Crespo-Bojorque and Toro, 2015; de la Mora, Nespor and Toro, 2013). In the current study, thirty-two rats were trained to distinguish between isochronous and non-isochronous beats at four different tempi (Inter Onset Interval of 300, 400, 500 and 600 ms). Forty sequences of twelve pure tones were presented to the rats in each session, and those sequences that were isochronous were reinforced with pellets. For each reinforced isochronous stimulus there was a unique non-isochronous stimulus lasting the same time and comprising the same number of beats in an irregular pseudorandom pattern (see Figure 1). The

deviant stimuli were never reinforced. Each standard-deviant couple had always the same temporal lapse between the first beat of the sequence and the last one.



Figure 1. An example of (a) an isochronous stimulus and (b) its non-isochronous counterpart.

After the training sessions, rats were tested three times with ten isochronous stimuli at two new tempi (IOI of 350 and 550 ms) and their non-isochronous counterparts. During the tests sessions, rats never received reward for the new isochronous and non-isochronous stimuli. A 2-way repeated measures ANOVA with the within-factors Test (T1, T2, T3) and Stimulus (Standard, Deviant) revealed two significant main effects: the nose-poking behavior of rats increased at each test and was significantly higher for the standard stimuli. In other words, rats insisted more on receiving the pellets over time and succeeded in distinguishing the isochronous beats from the non-isochronous ones.

This first approach to beat perception suggests that the perceptual timing mechanisms underlying isochrony detection are not limited to vocal learners. This is in line with studies on regularity detection in birds (Spierings and ten Cate, 2016; ten Cate, Spierings, Hubert and Honing, 2016) and the neural responses to the beat found in anaesthetized gerbils (Rajendran, Harper, Garcia-Lazaro, Lesica and Schnupp, 2017). Next studies will focus on more complex metrical rhythms, similar to those used with the mimicking starlings (Hulse, Humpal and Cynx, 1984), to see whether rats are able to deal with beat perception and meter induction beyond the level of isochrony.

Acknowledgements

This work was supported by the European Research Council (ERC) Starting Grant agreement n.312519 and by the Spanish Ministerio de Economía y Competitividad (MEC) FPI grant BES-2014-070547.

References

- Cook, P., Rouse, A., Wilson, M., & Reichmuth, C. (2013). A California sea lion (Zalophus californianus) can keep the beat: Motor entrainment to rhythmic auditory stimuli in a non vocal mimic. *Journal of Comparative Psychology*, *127*(4), 412.
- Crespo-Bojorque, P., & Toro, J. M. (2015). The use of interval ratios in consonance perception by rats (Rattus norvegicus) and humans (Homo sapiens). *Journal of Comparative Psychology*, *129*(1), 42.
- Fitch, W. T. (2006). The biology and evolution of music: A comparative perspective. *Cognition*, 100(1), 173-215.
- Fitch, W. T. (2013). Rhythmic cognition in humans and animals: distinguishing meter and pulse perception. *Frontiers in systems neuroscience*, 7.
- Hulse, S. H., Humpal, J., & Cynx, J. (1984). Discrimination and generalization of rhythmic and arrhythmic sound patterns by European starlings (Sturnus vulgaris). *Music Perception: An Interdisciplinary Journal*, 1(4), 442-464.
- Hattori, Y., Tomonaga, M., & Matsuzawa, T. (2013). Spontaneous synchronized tapping to an auditory rhythm in a chimpanzee. *Scientific reports*, *3*, 1566.
- Jarvis, E. D. (2006). Selection for and against vocal learning in birds and mammals. *Ornithological Science*, 5(1), 5-14.
- Jarvis, E. D. (2007). Neural systems for vocal learning in birds and humans: a synopsis. *Journal of Ornithology*, *148*(1), 35-44.
- Marler, P. (1970). Birdsong and speech development: Could there be parallels? There may be basic rules governing vocal learning to which many species conform, including man. *American scientist*, 58(6), 669-673.
- Marler, P. (1976). An ethological theory of the origin of vocal learning. *Annals* of the New York Academy of Sciences, 280(1), 386-395.
- Merchant, H., & Honing, H. (2013). Are non-human primates capable of rhythmic entrainment? Evidence for the gradual audiomotor evolution hypothesis. *Frontiers in Neuroscience*, *7*.
- De la Mora, D., Nespor, M., & Toro, J. M. (2013). Do humans and nonhuman animals share the grouping principles of the iambic-trochaic law?. Attention, Perception, & Psychophysics, 75(1), 92-100.
- Patel, A. D. (2006). Musical rhythm, linguistic rhythm, and human evolution. *Music Perception: An Interdisciplinary Journal*, 24(1), 99-104.
- Patel, A. D., Iversen, J. R., Bregman, M. R., & Schulz, I. (2009). Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Current biology*, 19(10), 827-830.

Patel, A. D. (2010). Music, language, and the brain. Oxford university press.

- Ravignani, A., & Madison, G. (2017). The paradox of isochrony in the evolution of human rhythm. *Frontiers in psychology*, 8.
- Ravignani, A., Gingras, B., Asano, R., Sonnweber, R., Matellán, V., & Fitch, W. T. (2013). The Evolution of Rhythmic Cognition: New Perspectives and Technologies in Comparative Research. In the *Proceedings of the 35th Annual Conference of the Cognitive Science Society*, Berlin, Germany, 31 July-3 August.
- Rajendran, V. G., Harper, N. S., Garcia-Lazaro, J. A., Lesica, N. A., & Schnupp, J. W. (2017, November). Midbrain adaptation may set the stage for the perception of musical beat. *Proceedings of the Royal Society B*, Vol. 284, No. 1866, p. 20171455. The Royal Society.
- Schachner, A., Brady, T. F., Pepperberg, I. M., & Hauser, M. D. (2009). Spontaneous motor entrainment to music in multiple vocal mimicking species. *Current Biology*, 19(10), 831-836.
- Spierings, M. J., & ten Cate, C. (2016). Zebra finches as a model species to understand the roots of rhythm. *Frontiers in neuroscience*, 10.
- ten Cate, C., Spierings, M., Hubert, J., & Honing, H. (2016). Can birds perceive rhythmic patterns? A review and experiments on a songbird and a parrot species. *Frontiers in psychology*, 7.

HOW DOMAIN-SPECIFIC IS MERGE?

DANIEL COOK^{*1}, MAURICIO J.D. MARTINS^{1,2}, ARNO VILLRINGER^{1,2}

*Corresponding Author: Dan Cook, cookdj0128@gmail.com
¹ Berlin School of Mind and Brain, Humboldt Universität zu Berlin, Berlin, Germany
² Neurology Department, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

1. Introduction

The Minimalist Program (MP) claims that Merge is the core cognitive computation responsible for the recursive hierarchies found in human language and thought (Chomsky, 2017a; Chomsky, 2017b; Hornstein, 2017). Additionally, it claims that Merge is unique to the Language Faculty (Berwick & Chomsky, 2016; Fitch et al., 2005; Hauser et al., 2002; Hornstein & Pietroski, 2009; Poeppel, 2017).

Recently, this language-uniqueness view was tested (Martins et al., 2015a) by asking participants to complete a recursive rule-based visual categorization task while performing a phonological working memory task. The verbal task did not interfere with the visual, suggesting that the representation of recursion in vision is not dependent on language domain-specific resources. This study can be criticized, however, for using a verbal task that is not linguistically domain-specific for the MP. More specifically, MP researchers have recently claimed that Merge consists of a linguistically domain-specific cognitive process of (1) lexicalization and (2) concatenation (Zaccarella & Friederici, 2016; Hornstein, 2017; Hornstein & Pietroski, 2009). Therefore, in the current study, we operationalize this critique and test the hypothesis that classifying images as well-formed continuations of a recursive rule relies on Merge as defined by the MP. To do so, we used the Visual Recursion Task in a dual-task paradigm that included interference conditions tapping specifically into components (1) and (2) of Merge.

In the Visual Recursion Task (VRT) (Martins, 2015b), participants are shown three successive iterations of a fractal. After the three images appear, two images are shown simultaneously and participants must select the image that best

corresponds to the correct continuation of the rule established by the first three images. In our study, 48 participants completed the VRT task with and without dual-task interference. The secondary tasks were: 1) A Semantic Fluency condition ('semantic') - tapping into lexical representations – in which participants received a basic category word (e.g. "animals") at the start of a VRT trial and named examples of this category for the duration of the trial; and 2) an Arithmetic Recursion condition ('math') - tapping into recursive concatenation – in which participants received a base value (N) at the start of a VRT trial and counted down backwards by three (N_i=N_{i-1}-3) for the duration of the trial. To compare any specific effects of dual-task interference on the VRT performance, in a second experiment, a new group of participants (N=48) completed a control task, he Embedded Iteration Task (EIT), a non-recursive iterative rule-based visual task (Martins, 2015b), with the same procedure and dual-task interference conditions as Experiment 1.

2. Analysis & Results

We analyzed accuracy and response times using a Generalized Linear Mixed Model in R (R Core Team, 2013; Bates & Sarkar, 2006). Subject ID was specified as a random factor to control for repeat measures across fixed factors. We performed model selection by adding factors in a stepwise procedure and comparing log likelihoods. In our final model, VRT trials in the 'math' and 'semantic' dual task conditions were associated with significantly higher error rates (p<.05; p<.001) and shorter response times (p<.05; p<.001) compared to trials without dual-task conditions.

Similar to Experiment 1, we analyzed accuracy and response times in the EIT experiment using a Generalized Linear Mixed Model in R. Compared to the EIT task without interference, the 'math' and 'semantic' dual task conditions were associated with significantly higher errors rates (p<.001; p<.001) and longer response times (p<.001; p<.001).

Lastly, we combined the VRT and EIT data to investigate the differential effects of the dual-task interference conditions on tasks. The effect of 'math' and 'semantic' interference in the VRT was associated with significantly *lower* error rates and *shorter* response times compared to the same categories of effects in the EIT (p<.001; p<.001).

3. Conclusion

Our results suggest that the ability to categorize visual recursion is dependent on some of the abilities enabling Semantic Fluency and Arithmetic Recursion but less so than the ability to interpolate fixed positions within a visual sequence (as required by the EIT). This suggests that subcomponents of Merge may be used in the generation of well-formed visual structures following rules. However, the rules may not only be used specifically by recursive operations thought to generate an unlimited amount of novel hierarchies of unbounded depth.

References

- Bates, D. and Sarkar, D. (2006), lme4: Linear Mixed-Effects Models Using S4 Classes, URL http://CRAN.R-project.org, R package version 0.99875-8.
- Berwick, R. C., & Chomsky, N. (2016). Why only us: Language and evolution. MIT press.
- Chomsky, N. (2017a). The language capacity: architecture and evolution. Psychonomic Bulletin & Review, 24(1), 200-203.
- Chomsky, N. (2017b). Language architecture and its import for evolution. Neuro. Biobehavioral Reviews. http://10.1016/j.neubiorev.2017.01.053
- Fitch, T., Hauser, M. & Chomsky, N. (2005). The evolution of the language faculty: clarifications and implications. Cognition. 97(2), 179–210.
- Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? Science, 298(5598), 1569-1579.
- Hornstein, N. (2017). The Minimalist Program After 25 Years. Annual Review of Linguistics, (0). https://doi.org/10.1146/annurev-linguistics-011817-045452
- Hornstein, N., & Pietroski, P. (2009). Basic operations: Minimal syntaxsemantics. Catalan Journal of Linguistics, 8, 113-139.
- Martins, M. D. J. D., Muršič, Z., Oh, J., & Fitch, W. T. (2015a). Representing visual recursion does not require verbal or motor resources. Cognitive Psychology, 77, 20-41.
- Martins, M. D., Martins, I. P., & Fitch, W. T. (2015b). A novel approach to investigate recursion and iteration in visual hierarchical processing. Behavior Research Methods, 48(4), 1421-1442.
- Poeppel, D. (2017). The Influence of Chomsky on the Neuroscience of Language. In J. McGilvray (Ed.), The Cambridge Companion to Chomsky (pp. 153-154). Cambridge: Cambridge University Press. doi:10.1017/9781316716694.008
- R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.
- Zaccarella, E., & Friederici, A. D. (2016). The neurobiological nature of syntactic hierarchies. Neuro. & Biobehav. Rev.. doi:10.1016/j.neubiorev.2016.07.038

STATISTICAL LEARNING AND LANGUAGE (IN SPITE OF ARBITRARINESS)

DAVIDE CREPALDI*1, SIMONA AMENTA2, and MARCO MARELLI3

*Corresponding Author: davide.crepaldi@sissa.it ¹Neuroscience Area, SISSA, Trieste, Italy ²Department of Psychology, University of Gent, Gent, Belgium ³Department of Psychology, University of Milano Bicocca, Milano, Italy

1. Introduction

It has long been known that the relationship between form and meaning is generally arbitrary in human languages, that is, forms have no inherent relationships with their meanings (Hockett, 1963). However, it has also been shown that the cumulative cultural evolution of languages does introduce regularities in form-tomeaning mapping (Kirby, Cornish, & Smith, 2008), and that systematicity in this mapping helps learnability, at least in terms of word categorization (Monaghan, Christiansen, & Fitneva, 2011). One apparent end product of this structureoriented "invisible hand" is linguistic morphology—families of words emerge whose relationship in form predicts their relationship in meaning (e.g., DEAL and DEALER, HUNT and HUNTER). Here we complement this evolutionary evidence with data from Cognitive Neuroscience, showing that the brain codes for these form-to-meaning regularities in a probabilistic way, and uses this information as we process words, either in isolation or embedded in sentence context.

2. Methods

As a test case, we considered the construct developed by Marelli, Amenta, and Crepaldi (2015), *Orthography-to-Semantics Consistency (OSC)*. This is a frequency-weighted average of meaning similarity between any given stem (e.g., DIAL) and all words that include that stem in their orthography (e.g., DIALECT, DIALLED, DIALS, DIALLING, DIALOG, DIALYSIS). Formally:

$$OSC(t) = \frac{\sum_{j=1}^{k} f_{r_x} \cos(\vec{t}, \vec{r_x})}{\sum_{j=1}^{k} f_{r_x}}$$
(1)

where t is a stem, f_{r_x} is the frequency of its k orthographic relatives r_x , and \vec{t} and $\vec{r_x}$ are vectorial representations of meaning as extracted from a distributional semantic model (Marelli et al., 2015).

Essentially, OSC tracks how strongly form similarity correlates with meaning similarity.

This measure was shown to predict word identification times in a large psycholinguistic database (Marelli et al., 2015; Balota et al., 2007). Here we test *OSC* against a behavioural index of sensitivity to form-meaning regularities, i.e., morphological priming; and against neurophysiological data (*Event-Related Potentials, ERP*) collected during natural sentence reading (Frank, Otten, Galli, & Vigliocco, 2015).

3. Results

OSC turns out to qualify morphological priming, in such a way that higher frequency primes have more impact on target processing, either strengthening priming, if they are indeed related to the target (e.g., corns–CORN); or weakening it, if they are not (e.g., corner-CORN).

Brain electrophysiology (the LAN and N400 components) is also shown to be modulated by OSC during sentence reading, at least in words that aren't easily predictable given the available sentence context.

4. Discussion

From a Cognitive Neuroscience point of view, these results suggest some rethinking of linguistic morphology. Rather than a discrete set of operations over a finite set of well-defined mental representations, morphology can (should?) be seen more generally as part of a form-to-meaning mapping effort carried out by the brain, on the basis of probabilistic knowledge that is accumulated through linguistic experience.

From an evolutionary perspective, these data show that the brain takes advantage of regularities in form-to-meaning mapping in language, thus establishing a psychological/neuroscience counterpart to the progressive emergence of structure through iterated learning. Of course, we don't know whether this phenomenon is an evolutionary reaction of the brain to the independently-triggered emergence of structure in language; or a core feature of the human cognitive machinery, which independently contributed itself to the emergence of structured form-to-meaning mapping. Interestingly, these data also link language learning/evolution/processing to general-purpose cognitive and brain operations (Ellison, 2013).

Acknowledgements

This work is supported in part by an ERC Starting Grant awarded to Davide Crepaldi (679010) and an FWO grant awarded to Marco Marelli, Marc Brysbaert and Simona Amenta (FWO.OPR.2017.0014.01).

References

- Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. I., Kessler, B., Loftis, B., Neely, J., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, 39, 445–459.
- Ellison, T. M. (2013). Categorisation as topographic mapping between uncorrelated spaces. In *Algorithmic Probability and Friends. Bayesian Prediction and Artificial Intelligence* (p. 131-141). Springer.
- Frank, S. L., Otten, L. J., Galli, G., & Vigliocco, G. (2015). The ERP response to the amount of information conveyed by words in sentences. *Brain and Language*, 140, 1–11.
- Hockett, C. F. (1963). The problem of universals in language. In J. H. Greenberg (Ed.), *Universals of Language* (pp. 1–29). Cambridge, Massachusetts: MIT Press.
- Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences, USA*, 105, 10681–10686.
- Marelli, M., Amenta, S., & Crepaldi, D. (2015). Semantic transparency in free stems: The effect of Orthography–Semantics Consistency on word recognition. *Quarterly Journal of Experimental Psychology*, 68, 1571–1583.
- Monaghan, P., Christiansen, M. H., & Fitneva, S. (2011). Balancing arbitrariness and systematicity in language evolution. In *The Evolution of Language* (p. 465-466). World Scientific.

ALIEN SYMBOLS FOR ALIEN LANGUAGE: ITERATED LEARNING IN A UNIQUE, NOVEL SIGNAL SPACE

CHRISTINE CUSKLEY^{*1}

*Corresponding Author: ccuskley@gmail.com ¹Centre for Language Evolution, The University of Edinburgh, UK

Studies in iterated learning of artificial 'alien' languages show that structure in language can arise from the process of cultural transmission (Kirby, Cornish, & Smith, 2008). Structure can accumulate not only in iterated artificial language learning (ALL) tasks, but also in iterated sequence learning tasks which are not explicitly linguistic (e.g., they contain no meanings, and do not use linguistic stimuli). Meaningless sequences of colours become more structured and more learnable over time (Cornish, Smith, & Kirby, 2013), reproduction of rhythmic sequences results in increased structure, learnability, and emergence of 'musical universals' (Ravignani, Delgado, & Kirby, 2016), and transmission of slide whistles can result increased structure (Verhoef, Kirby, & Boer, 2014). Where signals are mapped onto meanings, the relationships between these two spaces can have significant effects on the resulting structure (Little, Erylmaz, & Boer, 2017). While the signals and modalities with which iterated learning studies are conducted have expanded considerably in recent years, how properties of signal spaces interact with biases of learners in the emergence of structure is relatively under-studied.

To further investigate the effect of signal space on the evolution of structure and learnability, we used a completely novel set of graphical symbols called Ferros: a set of 137 abstract graphemes created using ferrofluid ink and visually unlike e.g., Roman orthography: (2) (2) . To articulate in Ferro, participants move in a virtual two dimensional palette which produces a different symbol depending on their location within a square. The space is structured along two axes corresponding to the number of contours and nodes in each Ferro. Much like the phonetic space of consonants, distance between Ferros encodes similarity, but the forms are discrete. Ferros are signals which are entirely foreign to participants, who have to learn not only to use the apparatus which produces Ferros, but also what the relevant features of Ferros are.

The Ferro palette was used in an iterated sequence copying task involving hundreds of participants: each participant had three minutes to copy as many 3-character Ferro sequences as they could. Each sequence faded slowly after presen-

tation, providing an additional pressure for quick reproduction. Participants were confined to producing three characters (i.e., sequences could not vary in length), and were provided with feedback on accuracy at each trial.

The initial sequences were pseudo-random, consisting of graphemes with either large articulation spaces in the palette or small, specific articulation spaces. Overall, 672 sequences were produced across 76 chains, each chain being between 5 and 10 generations. Reproduction error was measured as mean Euclidean distance in the palette from the target sequence to the produced sequence, and examined using a linear mixed effects model with generation and articulation space size as fixed effects.

Inclusion of generation and articulation space size improved the model fit significantly over a null model ($\chi^2 = 4.53$, p = 0.03). Error was higher overall in chains which started with small articulation spaces ($\beta = 0.1$, SE= 0.04, t = 2.5), indicating that participants found these harder to reproduce. Error decreased over generations in the small articulation space condition ($\beta = -0.018$, SE = 0.006, t = -3.035), but there was no such effect of generation for sequences which started with large articulation spaces.

Error decreased despite an overall retention of high sequence entropy: in other words, most sequences retained three different characters throughout a chain. The size of the area of the articulation space of a particular Ferro grapheme was significantly correlated with copying accuracy, and sequences in small articulation chains moved towards the larger articulation spaces over time, suggesting that larger articulation spaces make for greater ease of articulation.

The current study shows transmission can lead to increased learnability even in a completely novel signal system that participants have to learn from scratch in the process of transmission. In this case, the biases which drive increased learnability seem to derive primarily from the shape of the signal space: articulation area of a particular Ferro correlates significantly with its learnability (operationalised as reproduction error), and sequences gravitated towards larger articulation spaces over 'time'. The shape of the space was perhaps especially influential in the current task because participants had no prior experience with Ferros, and so had fewer (or less influential) prior biases relative to other iterated learning tasks, particularly those that utilise non-words and leverage existing linguistic knowledge. This has potential implications for the co-evolution of language and speech: constraints of the signal space may be more influential in early stages of emergence as 'speakers' are not only reproducing utterances, but also learning how to manipulate a novel signal space. This study shows that the Ferro palette has broad applications in language evolution research as a system of truly 'alien' symbols to study 'alien' language learning, emergence, and evolution.
Acknowledgements

This research was supported by British Academy Postdoctoral Fellowship grant no. pf150065. The author thanks Sean Roberts, Stefan Hartman, Michael Pleyer, and Hannah Little for valuable input with on the interface design, implementation, and assistance with translation.

- Cornish, H., Smith, K., & Kirby, S. (2013). Systems from sequences: an iterated learning account of the emergence of systematic structure in a nonlinguistic task. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th annual conference of the cognitive science society* (p. 340-345). Austin, TX: Cognitive Science Society.
- Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*, 105(31), 10681-10686.
- Little, H., Erylmaz, K., & Boer, B. de. (2017). Signal dimensionality and the emergence of combinatorial structure. *Cognition*, *168*, 1 15.
- Ravignani, A., Delgado, T., & Kirby, S. (2016). Musical evolution in the lab exhibits rhythmic universals. *Nature Human Behvaiour*, *1*.
- Verhoef, T., Kirby, S., & Boer, B. de. (2014). Emergence of combinatorial structure and economy through iterated learning with continuous acoustic signals. *Journal of Phonetics*, 43, 57 - 68.

CULTURAL TRANSMISSION OF MELODIC AND RHYTHMIC UNIVERSALS: 4 EXPERIMENTS AND A MODEL

TANIA DELGADO¹, ANDREA RAVIGNANI^{*2,3,4}, TESSA VERHOEF^{#5,6}, BILL THOMPSON^{2,4}, THOMAS GROSSI⁷, SIMON KIRBY⁷

*Corresponding Author: andrea.ravignani@gmail.com #Corresponding Author: tessauva@gmail.com

¹Department of Cognitive Science, University of California, San Diego, La Jolla, USA ²Artificial Intelligence Lab, Vrije Universiteit Brussel, Brussels, Belgium ³Research Department, Sealcentre Pieterburen, Pieterburen, Netherlands

⁴Language and Cognition Department, Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands

⁵Leiden Institute of Advanced Computer Science (LIACS), Leiden University, Leiden, The Netherlands

⁶Center for Research in Language, University of California, San Diego, La Jolla, USA ⁷Centre for Language Evolution, School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, UK.

Humans are well-versed at processing sequences. In the lab, when confronted with language or other culturally-transmitted systems, humans introduce and amplify structural regularities making the systems easier to learn (Kirby, Griffiths, & Smith, 2014; Kirby, 2017). Is the cultural emergence of fine-grained regularities a prerogative of language alone? Can cultural transmission explain universals in musical structure (Savage et al., 2015; Trehub, 2015) as it explains 'linguistic universals' (in the Greenbergian sense)? We tackled these questions in the lab and in-silico by adopting an iterated learning paradigm. Two experiments addressed the evolution of *rhythmic* structure (Ravignani et al., 2017), and two the evolution of *melodic* structure (using, among others, data from Verhoef, 2012; Verhoef et al., 2014). Depending on the experiment, participants were given a slide whistle or an electronic drum kit, and were asked to imitate a sound sequence to the best of their abilities. The output of one generation of participants became the input of the next generation.

Across experiments we varied conditions known to affect transmission chains, such as immediate vs. delayed recall, within-participant vs. between-participant transmission, etc. We also introduce a probabilistic model for the latent structures underpinning rhythmic sequences, alongside a psychologically plausible posterior sampling algorithm (Ravignani et al., 2017). This allows us to obtain approximate structural descriptions of rhythmic patterns across conditions and generations. The initially random experimental stimuli became more musical. In particular, most regularities emerging from our transmission chains are statistical universals of world music (Savage et al., 2015; Fitch, 2017). Both whistles and drumming patterns became more compressible, measured in terms of decreasing entropy, and easier to learn. We found reuse of a small set of basic building blocks in the emerging systems, resulting in more predictable sequences of sounds (Ravignani, 2017). This corresponds with a musical universal, namely the repetition of melodic and rhythmic phrases in music. This reuse of elements is accompanied by a transition from continuous to discrete use of pitch contours (showing convergence to another universal; Ravignani & Verhoef, 2017). Participants produced sequences containing melodic and rhythmic patterns, i.e. musical motifs. We also see a gradual increase in mirrored elements, which suggests that the emerging melodies contain arched contours (a common musical universal). In addition, drumming sequences became more isochronous (Ravignani & Madison, 2017), and composed of few (categorically distributed) alternating inter-beat intervals, related by small integer ratios. Patterns transformed by between-participants transmission show similar properties to those emerging from within-participant transmission (i.e. self learning; Ravignani et al., 2017). Other melodic universals, such as the length of phrases, frequency intervals of melody contours, and organization of scales can also be measured in this data set and are currently being tested. Analysis of the probabilistic model supports these insights, suggesting that later generations show increased re-use of prototypical building blocks both within and across individual sequences (Ravignani et al., 2017). The emergence of musical structure via cultural transmission: (a) does not require semantics or learning language-like behaviours; (b) operates similarly across domains of human cognition; (c) explains characteristics of music appearing as statistical universals around the world (Savage et al., 2015).

Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant

agreement No. 665501 with the research Foundation Flanders (FWO) (Pegasus2 Marie Curie fellowship 12N5517N awarded to A.R.), a visiting fellowship in Language Evolution from the Max Planck Society (awarded to A.R.), and ERC Grant [283435] ABACUS (awarded to B.d.B.).

- Fitch, W. T. (2017). Cultural evolution: Lab-cultured musical universals. *Nature Human Behaviour*, *1*, 0018.
- Kirby, S., Griffiths, T., & Smith, K. (2014). Iterated learning and the evolution of language. *Current opinion in neurobiology*, 28, 108-114.
- Kirby, S. (2017). Culture and biology in the origins of linguistic structure. *Psychonomic bulletin & review*, 24(1), 118-137.
- Ravignani, A., Delgado, T., & Kirby, S. (2016). Musical evolution in the lab exhibits rhythmic universals. *Nature Human Behaviour*, *1*, 0007.
- Ravignani, A. (2017). Visualizing and interpreting rhythmic patterns using phase space plots. *Music Perception: An Interdisciplinary Journal*, 34(5), 557-568.
- Ravignani, A., Thompson, B., Grossi, T., Delgado, T., & Kirby, S. (2017). Evolving building blocks of rhythm: How human cognition creates music via cultural transmission. *bioRxiv*, 198390.
- Ravignani, A., & Madison, G. (2017). The paradox of isochrony in the evolution of human rhythm. *Frontiers in Psychology*, 8.
- Ravignani, A., & Verhoef, T. (2017). Which melodic universals emerge from repeated signaling games? (No. e3366v1). *PeerJ Preprints*.
- Savage, P. E., Brown, S., Sakai, E., & Currie, T. E. (2015). Statistical universals reveal the structures and functions of human music. *Proceedings of the National Academy of Sciences*, 112(29), 8987-8992.
- Trehub, S. E. (2015). Cross-cultural convergence of musical features. Proceedings of the National Academy of Sciences, 112(29), 8809-8810.
- Verhoef, T. (2012) The origins of duality of patterning in artificial whistled languages. *Language and Cognition* 4(4), 357-380.
- Verhoef, T., Kirby, S. & de Boer, B. (2014) Emergence of combinatorial structure and economy through iterated learning with continuous acoustic signals. *Journal of Phonetics* 43C, pp. 57-68

PHYLOGENY IN PHONOLOGY: TAI SOUND SYSTEMS ENCODE THEIR EVOLUTIONARY HISTORY

RIKKER DOCKUM *1

*Corresponding Author: rikker.dockum@yale.edu ¹Department of Linguistics, Yale University, New Haven, CT, USA

It is well established that the phonology of modern languages encodes useful data for studying the history of languages and their genealogical relationships to one another. The study of sound change is a foundation of traditional historical linguistics, particularly the linguistic comparative method. However, the utility of phonological systems for computational phylogenetic study has been more restricted. Phonology has typically been the means to an end: traditional comparative analysis enables coding a lexical dataset for cognacy. But once coded in this manner, the specifics of the phonology no longer factor into the quantitative analysis. This paper uses Tai languages to demonstrate that the facts of phonological systems themselves encode phylogenetic signal directly, suitable for use in quantitative historical analysis.

This study extends findings in previous by e.g. Macklin-Cordes (2015) and Macklin-Cordes and Round (2015). In this study, lexical data from Hudak (2008) comprising 1,159 cognate sets covering 20 Tai lects were used to generate datasets of two broad types: (1) binary data, encoding the presence or absence of phones and biphones in the lexicon of each language under examination, and (2) continuous data, which encodes phone frequency and biphone Markov chain transition probability (Ching & Ng 2006) in each language. Together these form a phonotactic profile for each language. Only those phones and biphones that exhibit variation in the overall dataset were included in the statistical analyses. The hypothesis underlying this type of data is that the more closely related two languages are to one another, the more similar their phonotactic profile will be. The phylogenetic tree required by these statistical tests was adapted from Chamberlain (1975).

Three tests are used to examine phylogenetic signal in the datasets described: D test (Fritz & Purvis 2010), Blomberg's K (Blomberg et al 2003), and *NeighborNet* (Bryant & Moulton 2004), which also involves the calculation of a delta-score (Holland et al 2002), and the mean Q-residual (Gray et al 2010). The Density plots of D and K values for phones and biphones are in Figures 1 to 4.







Figures 3 and 4. Density plots of K for phoneme frequency and biphone transition probability.

The results of these statistical tests confirm findings in Macklin-Cordes (2015) of phylogenetic signal in the "high-resolution" phonotactic profiles of languages. They also extend those findings by detecting strong signal in binary presence/absence data where that study was unable to do so. The results of the present study are of interest to linguists generally in the ongoing work of developing and testing phylogenetic methods of linguistic analysis. While the relative difficulty of using the traditional linguistic comparative method with Australian languages makes phylogenetic tools especially attractive and useful, the demonstrated results with the Tai data also shows the potential utility of these methods in other language families where traditional methods already have traction. This thus serves as a model for the application of these tests to language families and geographical regions in need of improved language classification throughout the world.

- Blomberg, S. P., T. Garland, & A. R. Ives. (2003). Testing for Phylogenetic Signal in Comparative Data: Behavioral Traits Are More Labile. *Evolution* 57(4): 717–745.
- Bryant, D., & Moulton, V. (2004). Neighbor-Net: An Agglomerative Method for the Construction of Phylogenetic Networks. *Molecular biology and evolution* 21(2): 255–265.
- Fritz, S. A., & Purvis, A. (2010). Selectivity in Mammalian Extinction Risk and Threat Types: A New Measure of Phylogenetic Strength in Binary Traits. *Conservation Biology* 24(4): 1042–1051.
- Gray, R. D., Bryant, D., & Greenhill, S. J. (2010). On the shape and fabric of human history. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 365(1559): 3923-3933.
- Holland, B. R., Huber, K. T., Dress, A., & Moulton, V. (2002). Delta Plots: A Tool for Analyzing Phylogenetic Distance Data. *Molecular biology and evolution* 19(12): 2051–2059.
- Hudak, Thomas John. (2008). William J. Gedney's Comparative Tai Sourcebook. Oceanic Linguistics Special Publication No. 34. Honolulu: University of Hawai'i Press.
- Macklin-Cordes, Jayden. (2015). *Phylogeny and Phonotactics: Quantifying Historical Signal in Sequences of Sound*. B.A. thesis. University of Queensland.
- Macklin-Cordes, Jayden & Round, Erich. (2015). High-definition phonotactics reflect linguistic pasts. Proceedings of the 6th Conference on Quantitative Investigations in Theoretical Linguistics, Tubingen, Germany.

METAPHORICAL EXTENSION AND THE EVOLUTION OF CONFIGURATIONALITY

T. Mark Ellison^{*1} and Uta Reinöhl²

*Corresponding Author: m.ellison@anu.edu.au
¹Wellsprings/CoEDL, Australian National University, Canberra, Australia
²Institut für Linguistik, Universität zu Köln, Köln, Germany

Syntactic systems fall along a spectrum of *configurationality*. At one end, *less-configurational* languages arrange their sentences by information structure while at the *more-configurational* other end, word order is constrained by syntactic rules and requires the overt expression of specific words (e.g. adpositions requiring overt nominal arguments). Historical records show that some more-configurational languages, like English or Hindi, have developed from less-configurational (or *non-configurational* in the strict Australianist sense, Hale 1983) ancestral languages (Luraghi 2010, Reinöhl 2016). We present a computational model of the development of certain aspects of configurationality, drawing on work by Reinöhl (2016).

In studying the development of adpositions from nominals, Reinöhl (2016) discovered that metaphorical extension of relational terms leads to overt dependents becoming obligatory. Here, *metaphorical* means any usage where the relational term is combined with a semantically incompatible argument - incompatible under literal interpretation. For instance, the literal meaning of the relational term *middle* is only coherent with arguments that contain an actual, semantically retrievable centre (= roughly equidistant from the container's spatial and/or temporal boundaries) as in *in the middle of the room* or *in the middle of the day*. Besides this literal usage, *middle* can be used metaphorically when there is no well-defined centre. In the example A moment ago everything was OK. Suddenly, we are in the middle of a crisis, the word middle does not refer to the temporal centre of the crisis, in equidistance from its edges: the crisis may last for a week, but within 5 minutes we are in its middle.

When a relational term is used in a non-literal, metaphorical way as just illustrated, the target domain (e.g. *the crisis*) shows a strong bias for being overtly encoded in the same clause. In contrast, literal usages of *middle* need not specify a possessor, so long as it is already primed, as in *The players arrived at the soccer field, the referee already standing in the middle*. A British National Corpus (BNC) search of relevant constructions involving *middle* shows a bal-

anced mix of possessed and unpossessed forms with literal interpretations, but only exceptional unpossessed forms where the interpretation is metaphorical.

Historical evidence from three millennia of Indo-Aryan history, stretching from Vedic Sanskrit to modern Hindi, shows the same phenomenon. In stage one (see ex. 1a), the Vedic Sanskrit noun $\exists \exists madhye$ in the locative case, meaning 'in the middle', is used in a basic spatial sense, i.e. referring to the middle point in spatial (or sometimes temporal) entities with an inherent centre, such as convex objects.

In the second stage (see ex. 1b), around two millennia later, *madhye* (now as *majjhe/-i*) occurs more often metaphorically, i.e. with entities that do not have inherent centres, and may not even be spatial or temporal. In these metaphorical usages, the possessor argument is always overt. In stage three (see ex. 1c), increasing metaphorical use over time leads to syntactic fixation of this construction. *Madhye* (as modern Hindi $m\tilde{e}$) has become semantically bleached, denoting simple containment (translatable as 'in'), and requires an explicit nominal dependent in all cases. This transition from a spatial noun to an adposition mirrors developments in numerous languages (Svorou 1988). The Indo-Aryan case study unveils how the dependents of adpositions may become obligatory due to incremental, semantically driven change.

1a) Vedic / literal use	1b) Apabhramśa / metaphor	1c) Hindi / semantic shift
atha madhya āghārayati	majjhi mahattarāṇa	vah pareśānī mẽ hai
now middle pour.3sg	middle of-best-ones	3sg trouble.obl.sg.f in be.3sg
now he pours (ghee) onto the middle (of the altar - omitted)	in the middle of (i.e. among) the best ones	s/he is in trouble

Hough et al. (2015) defines a set of desiderata for incremental dialogue systems, both as models of human cognitive processing and for natural language processing. One desireable feature is *monotonicity*: that all inferences true from the discourse before an input word is processed should still hold after it is processed. The processing of metaphors monotonically, however, leads to inconsistent representations, and thus wrong interpretations, unless the inconsistency is resolved by processing the metaphor - including a domain specifier - as a unit before integrating the new knowledge into the discourse representation. If metaphorical uses become sufficiently frequent, learners may infer that the argument is obligatory, while at the same time generalising the semantic sense of the term so that once metaphorical senses now become literal. We present simulations of obligatorification created by implementing one possible model of monotonic, incremental semantic parsing.

In summary, this poster describes how the process of metaphorical extension can lead to the obligatorification of arguments.

- Hale, K. (1983). Warlpiri and the grammar of non-configurational languages. *Natural Language & Linguistic Theory*, *1*(1), 5-47.
- Hough, J., Kennington, C., Schlangen, D., & Ginzburg, J. (2015). Incremental Semantics for Dialogue Processing: Requirements and a Comparison of Two Approaches. In Proceedings of the 11th International Conference on Computational Semantics (IWCS) 2015.
- Luraghi, S. (2010). The rise (and possible downfall) of configurationality. In S. Luraghi & V. Bubenik (Eds.), *The continuum companion to historical linguistics* (p. 212-229). London/New York: Continuum.
- Reinöhl, U. (2016). *Grammaticalization and the Rise of Configurationality in Indo-Aryan*. Oxford: Oxford University Press.
- Svorou, S. (1988). The experiential basis of the grammar of space. evidence from the languages of the world. (Unpublished doctoral dissertation). SUNY.

FROM COGNITION TO LANGUAGE CHANGE: CLOSE SIMULATION OF A PNG VILLAGE

T. Mark Ellison^{*1} and Eri Kashima¹

^{*}Corresponding Author: m.ellison@anu.edu.au ¹Wellsprings/CoEDL/CHL, Australian National University, Canberra, Australia

Research in psycholinguistics and experimental semiotics provide us with new understandings of human communicative processes. Agent-based simulations offer the enticing possibility of converting this understanding into explanations of language change (e.g. Stanford & Kenny 2013). Such simulations can only be accurate, however, if the interactions between agents pattern like interactions between language community members. In this talk, we present a one-person-to-one-agent simulation of a village in southern PNG. To maintain individual anonymity, we will call the village V, and their language L. The opportunities for such simulations are rare, and the only similar study (of which we know) is a model of a small Pacific island community (Hendery & Magee 2015).

While language L has closely related languages to its east and west, it has relatively little contact with unrelated languages. It has a small population base - less than a thousand speakers - but does not seem at immediate risk of being supplanted by English, Tok Pisin or Motu.

V is the smallest of the three villages which identify as L-speaking. It has 136 inhabitants living across 20 households. Life in the village revolves around gardening with the nuclear family, and gender-based activities such as fishing by women. Gardening occurs at hamlets which may be 30 minutes to two hours walk away from the village. These hamlets may be shared between related (or sometimes unrelated) households. During the non-wet seasons between August to December people often stay overnight at the hamlets in shelters built there. During the height of the wet season between January and April people are mostly confined to their houses in the village.

Of the 136 inhabitants in V, 100 are born after 1980. The gender split is 54% Male to 46% Female, however the percentage is skewed towards women in the older residents born before 1980 (53% Female, 47% Male). Of the 18 married women in the village, 5 have married in from non-L-speaking villages. The small population of the village and the completeness of the data we have on it allow us to build a close simulation of likely interactions.

We simulate language maintenance and innovation in V by combining de-

mographic and cognitive modelling. The demographic model combines features of actual persons: domicile proximity, age, first language, gender and gardening area, to determine the interaction likelihood of the corresponding agents. Age also impacts the strength of the egocentric bias (due to weight of past experience). Linguistic change occurs, if at all, as a result of interactions sampled from this likelihood distribution. In each interaction, one randomly selected participant is the speaker, another the hearer.

Tamariz, Ellison, Barr, and Fay (2014) develop a model of selection between variant representations based on earlier experimental semiotics studies. This model assumes that production choices reflect distributions of past experience, modified by two biases. These are the content bias - one variant being inherently more appealing, and an egocentrism bias - the priority agents give to repeating their past representations vs adopting representations which others have used. We adopt this model to account for selection between alternative linguistic forms in V during interactions.



Figure 1. The probability of using introduced variant over time for various levels of content bias (β) .

achieves some uptake.

We explore tentative evidence for a sound-change in progress in V, and what it can tell us about starting conditions and bias levels. In future work, we will integrate more accurate detailed conditioning of interaction likelihood on inhabitant attributes and daily routines.

One explanation of language change in exogamous communities is that inmarrying women are vectors for new variants; while not the only crosslinguistic contacts in the community, they are arguably the strongest. We explore this option in simulation: with in-married women initialised with two variants at equal likelihood, while everyone else in the village uses only a single variant. People are assumed to be egocentric in their option selection, the strength of this bias to be determined by fit-to-data. Figure 1 explores the impact of varying levels of attractiveness in the new variant. The more attractive the variant, the faster it spreads to the rest of the community. Interestingly, even without a content bias for the new variant, it still

- Hendery, R., & Magee, L. (2015). Geo-language games: An agent-based model of the role of terrain in language diversity. In *Digital humanities 2015*. Sydney. Retrieved from http://dh2015.org/abstracts/xml/ HENDERY_Rachel_Marion_Geo_Language_Games__An_Agen/
- Stanford, J. N., & Kenny, L. A. (2013, July). Revisiting transmission and diffusion: An agent-based model of vowel chain shifts across large communities. *Language Variation and Change*, 25(2), 119–153.
- Tamariz, M., Ellison, T. M., Barr, D. J., & Fay, N. (2014). Cultural selection drives the evolution of human communication systems. *Proceedings of the Royal Society B: Biological Sciences*, 281(1788), 20140488.

MEANING-REFINING ACOUSTIC VARIATION WITHIN THE INTERNAL STRUCTURE OF PIED BABBLER RECRUITMENT CRIES

SABRINA ENGESSER $^{\ast 1}$, AMANDA R. RIDLEY 2 , MARTA B. MANSER 3 , and SIMON W. TOWNSEND 1,4

*Corresponding Author: sabrina.engesser@uzh.ch
¹Department of Comparative Linguistics, University of Zurich, Zurich, Switzerland
²Centre for Evolutionary Biology, The University of Western Australia, Crawley, Australia
³Department of Evolutionary Biology and Environmental Studies, University of Zurich, Switzerland
⁴Department of Psychology, University of Warwick, Coventry, UK

Human language is a generative system that achieves its limitless productivity by assembling sounds into larger meaningful constructs (Chomsky, 1981; Hockett, 1960). Comparative work has demonstrated animals can also combine sounds in ways analogous to humans and this has shed important light on the prevalence of combinatorial capacities outside of humans (primates: Arnold & Zuberbühler, 2006; Ouattara, Lemasson, & Zuberbühler, 2009 / birds: Engesser et al., 2015; Engesser et al., 2016; Pepperberg, 2007; Suzuki et al., 2016 / see also: Hurford, 2007; Marler, 1977). However, further analogue examples are central in unveiling the distribution and diversity of combinatorial mechanisms in non-human communication systems, and to identify potential factors driving their emergence (Schlenker, Chemla, & Zuberbühler, 2016).

Here we present evidence for a novel form of combinatorial structuring in the vocal system of a highly social passerine with a fixed vocal repertoire, the southern pied babbler (*Turdoides bicolor*) (Ridley & Raihani, 2007). By studying a population of wild, but habituated, babblers we demonstrate that male babblers produce two variants of long and raucous, 'cry-like' structures, which generally appear to function to recruit group members during group travel. Using acoustic analyses, we show that both cry variants are similar in their

super-structure, but differ in their sub-structure. Preceded by a wind-up segment, they either grade into repetitions of A/single-note or AB/double-note motifs, with the same A-notes being shared across the two variants (Fig. 1). Behavioural observations on the natural occurrence of the two cry types in combination with playback experiments indicated that, consistent with similarities in their superstructure, both types function overall in recruiting group members during locomotion, but the internal A or AB sub-structure specifies the precise form of recruitment. Specifically, in response to A/single-note cries receivers were found to approach to the caller's announced location (i.e. tantamount to 'come to me'), while AB/double-note cries were associated with caller movement and with receivers following the caller over long distances (i.e. tantamount to 'come with me'). We suggest that the overall structure of the two cry variants likely conveys the same intention of the caller to recruit its group members, with the internal motif pattern refining the signal's functional specificity. Accordingly, the B note might represent an acoustic modifier altering or intensifying the A note's meaning.

We argue the pied babbler recruitment cry represents another intriguing example illustrating the variability of generative mechanisms outside of human language. Our work lends support to the hypothesis that combinatoriality emerged in species with constrained sound repertoires, whereby the assemblage of sounds into more distinctive structures might enhance signal discrimination, and hence increase communicative output (Arnold & Zuberbühler, 2008; Nowak & Krakauer, 1999). Ultimately, by unveiling potential conditions promoting the emergence of combinatorial capacities, such comparative data on non-human animals can provide valuable insights into the evolutionary progression of our own language system.



Figure 1. Spectrograms of a single-note (SN) and a double-note (DN) recruitment cry of one dominant male babbler. Capital letters denote the note type.

- Arnold, K., & Zuberbühler, K. (2006). Language evolution: Semantic combinations in primate calls. *Nature*, 441(7091), 303.
- Arnold, K., & Zuberbühler, K. (2008). Meaningful call combinations in a nonhuman primate. *Current Biology*, 18(5), R202-R203.
- Chomsky, N. (1981). Knowledge of Language: Its Elements and Origins. *Philosophical Transactions of The Royal Society B: Biological Sciences*, 295(1077), 223-234.
- Engesser, S., Crane, J. M., Savage, J. L., Russell, A. F., & Townsend, S. W. (2015). Experimental Evidence for Phonemic Contrasts in a Nonhuman Vocal System. *PloS Biology*, 13(6), e1002171.
- Engesser, S., Ridley, A. R., & Townsend, S. W. (2016). Meaningful call combinations and compositional processing in the southern pied babbler. *Proceedings of the National Academy of Sciences*, 113(21), 5976-5981.
- Hockett, C. F. (1960). The Origin of Speech. Scientific American, 203, 88-111.
- Hurford, J. (2007). The origins of meaning. Oxford: Oxford University Press.
- Marler, P. (1977). The structure of animal communication sounds. Recognition of complex acoustic signals: report of Dahlem workshop. Berlin: Abakon Verlagsgesellschaft.
- Nowak, M. A., & Krakauer, D. C. (1999). The evolution of language. Proceedings of the National Academy of Sciences, 96(14), 8028-8033.
- Ouattara, K., Lemasson, A., & Zuberbühler, K. (2009). Campbell's monkeys concatenate vocalizations into context-specific call sequences. *Proceedings* of the National Academy of Sciences, 106(51), 22026-22031.
- Pepperberg, I. M. (2007). Grey parrots do not always 'parrot': the roles of imitation and phonological awareness in the creation of new labels from existing vocalizations. *Language Sciences*, 29(1), 1-13.
- Ridley, A. R., & Raihani, N. J. (2007). Variable postfledging care in a cooperative bird: causes and consequences. *Behavioral Ecology*, 18(6), 994-1000.
- Schlenker, P., Chemla, E., & Zuberbühler, K. (2016). What Do Monkey Calls Mean? Trends in Cognitive Sciences, 20(12), 894-904.
- Suzuki, T. N., Wheatcroft, D., & Griesser, M. (2016). Experimental evidence for compositional syntax in bird calls. *Nature Communications*, 7, 10986.

STRUCTURAL CUES FOR SYMMETRY, ASYMMETRY AND NON-SYMMETRY IN CENTRAL TAURUS SIGN LANGUAGE

RABIA ERGIN*1, ANN SENGHAS2, RAY JACKENDOFF3, and LILA GLEITMAN4

^{*}Corresponding Author: Rabia.Ergin@mpi.nl ¹Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands ²Department of Psychology, Barnard College, New York City, USA ³Department of Brain and Cognitive Sciences, MIT, Cambridge, USA ⁴Department of Psychology, University of Pennsylvania, Philadelphia, USA

Central Taurus Sign Language (CTSL) is a village sign language used in three villages in an isolated area in south-central Turkey. This language emerged in the absence of a linguistic model within the last half-century as a result of recessive deafness in these communities, which have deaf populations of 4.8%, .6% and .5%. CTSL provides us with a novel vantage point into how a brandnew system emerges and develops because it is relatively young, still evolving, and the very first creators of this system are still alive today.

Previous research has shown that symmetrical and reciprocal predicates have specific and distinct syntactic properties in spoken languages (Gleitman, Gleitman, Miller & Ostrin, 1996) and in Nicaraguan Sign Language (NSL) (Flaherty, Goldin-Meadow, Senghas, Coppola & Gleitman, 2013; Flaherty, Goldin-Meadow, Senghas, Coppola & Gleitman, 2014). The present study investigated (i) whether CTSL has any distinctive structural cues marking symmetrical actions (e.g., shaking hands) and reciprocal actions (e.g., pushing each other); (ii) if it has, how these structural markers evolve over time across age cohorts of CTSL signers. We developed a controlled elicitation task in which deaf signers viewed 62 short clips and described them to a deaf/hearing addressee, who then picked the corresponding picture from an array of three pictures for comprehension check. The clips involved two characters performing actions that were symmetrical (e.g., shaking hands), reciprocal (e.g., punching each other), transitive (e.g., one punching the other), and intransitive (e.g., both punching towards the camera). Here we use symmetrical for events that are necessarily symmetric, reciprocal for events that are symmetric but can also appear in asymmetric/non-symmetric contexts, transitive for non-symmetric two- or three-argument events, and intransitive for non-symmetric one-argument events. Twelve signers, grouped into three successive age cohorts, participated $(M_{age}=42.2, age range: 17-55)$. We detected several devices that were candidate

structural cues, and compared them across cohorts as a measure of how CTSL has evolved over its 50-year existence.

Our findings, based on a total of 946 responses (n intransitive = 255, n transitive = 364, n reciprocal = 190, n symmetrical = 137), are as follows: (1) Body segmentation: Symmetrical and reciprocal actions come to be frequently marked by body segmentation -one side of the body is allocated for one of the characters and the other side is allocated for the other character- but not transitive actions (Figure 1). (2) *Mirroring*: Signing with both hands in a mirror-image configuration (cf. Flaherty et al, 2014) was often used for reciprocal and symmetrical actions, but not for intransitive and transitive actions (Figure 2). (3) Temporal sequencing: CTSL signers distribute information temporally across an utterance by sequentially signing each action performed by each character in the contexts involving intransitive, transitive and reciprocal actions, but not symmetrical actions. Temporal sequencing becomes more systematic in successive cohorts from CTSL-1 to CTSL-2 (χ^2 (1)= 43.4, p<0.0001) (Figure 3). (4) Perspective: Flaherty et al. (2013) found evidence for double perspective verb pairs in reversible transitive contexts in NSL -events being expressed both from the agent's and the patient's perspective. In contrast, here we find that CTSL prefers a single perspective in reversible transitive events, but a double perspective in reciprocal and symmetrical events, with an increasing tendency for systematicity in reciprocal events from CTSL-1 to CTSL-2 (χ^2 (1)= 48.4, p<0.0001) (Figure 4).

Briefly, our results indicate that, *first*, body segmentation and mirroring are strong tendencies as of CTSL-1 because signers makes use of the iconicity of the body, whereas temporal sequencing and perspective take time to invent as structural markers. Second, reciprocal and symmetrical actions differ from transitive actions in that the transitive actions do not use body segmentation and mirrored articulators, and, they are expressed from a single perspective, whereas reciprocal and symmetrical actions are body segmented, mirrored, and expressed from double perspectives. *Third*, reciprocal and symmetrical actions differ from plural intransitives by mirror-image configuration. Fourth, a reciprocal action differs from a symmetrical action in that the former is temporally sequenced whereas the latter is not. All in all, each action has its own combination of essential components to encode symmetry, asymmetry, and non-symmetry, and these components become more conventionalized across cohorts. Our findings provide further evidence for how a brand-new language converges on distinctive shared devices that differentiate between verb classes in an increasingly systematic way over generations of learners. More broadly speaking, the rapid cultural development of linguistic expressions that distinguish reciprocal and symmetrical actions from transitive actions and from each other suggests that a sensitivity to the semantics of these distinctions is present in the language-ready brain, and that this sensitivity is a product of biological evolution of the human linguistic and/or conceptual capacity.





Figure 3. Overall proportion of temp. seq across cohorts

Figure 4. Overall proportion of double. perspective across cohorts

Reciprocal

Symmetrical

- Flaherty, M., Goldin-Meadow, S., Senghas, A., Coppola, M. & Gleitman, L. (2014, April). Language from Gesture? Emergent Transitivity Marking in Nicaraguan Sign Language. Paper presented at the 10th Evolution of Language Conference, Vienna, Austria.
- Flaherty, M., Goldin-Meadow, S., Senghas, A., Coppola, M. & Gleitman, L (2013, November). Animacy and Verb Classes in Nicaraguan Sign Language. Poster presented at the 38th Annual Boston University Conference on Language Development, Boston, MA.
- Gleitman, L., Gleitman, H., Miller, C., & Ostrin, R. (1996). 'Similar' and similar concepts. Cognition, 58, 321-376.

DUAL-PROCESSING AND THE REPRESENTATIONAL HYPOTHESIS: ACCOUNTING FOR THE EMERGENCE OF LANGUAGE

ANDREW FEENEY

andrew.feeney@northumbria.ac.uk

Department of Humanities, Northumbria University, Newcastle upon Tyne, United Kingdom

There are well-rehearsed arguments, including those that appeal to aspects of psycholinguistic processing and language acquisition, that suggest that language has certain properties that distinguish it from domain general capacities. However, I argue that this position is based on a misreading of the properties, and relationship, of language and thought.

In the first instance there is a fundamental dichotomy: either humans think in the languages they speak, or in another, unarticulated system. I agree with Penn et al. that 'the adaptive advantages of being able to reason in a relational fashion have a certain primacy over the communicative function of language' (2008: 123). Language as a system of symbolic representation could not have evolved if conceptual categories did not already exist. The nature of such cognition is central to Dual-Processing Theory which is founded on a substantial body of evidence that suggests that modern humans are characterized by two fundamentally distinctive modes of reasoning (Eagleman, 2011; Evans, 2010). For while we share with nonhuman animals a similar system of fast, intuitive cognitive processing based on evolutionarily older brain systems - System 1 (Frankish, 2009) – we have also, at some point in the last 7.5 million years or more, evolved a capacity for systematic, computational thought that enables us to be logical and reflective and on occasions to override instinctive, associative reactions (System 2). The questions in relation to language evolution are: when did these developments occur and what is the relation to language?

108

Drawing on recent findings from ethology, evolutionary biology and paleontology it becomes clear that there were two significant periods in the evolution of human cognition corresponding to stages of punctuated equilibrium (Gould and Eldredge, 1993). The first of these followed a period, approximately two million years ago, in which there were a comparatively large number of changes to genes and genomic regions (especially in the Human Accelerated Region 1) with consequences for brain lateralization, connectivity and overall brain size (Kamm et al., 2013). This was associated with the first irrefutable evidence of changes to cognitive behavior resulting in sophisticated (mode 2) tool use, migration out of Africa and probable use of fire (Beyene et al, 2012; Wynn, 2012; Lynch and Granger, 2008). Such accomplishments are indicative of cooperation beyond that exhibited by any preceding hominins, and I agree with Levinson and Evans that 'there can be no doubt that premodern humans were talking' (2010: 2742). I suggest that the first of these linguistic hominins belonged to a species such as Homo erectus and were endowed with a protolanguage involving the simple concatenation of symbols, lacking the structural complexity that characterizes modern human language (Jackendoff, 2002; Tallerman, 2012).

The following period of a million years or so, although witness to some physical adaptations that benefited vocalization, was characterized by almost complete cultural stasis indicative of species with a cognitive capacity that lacked substantial creativity and was essentially still of type System 1. Until, around 500 thousand years ago, there was a second intense episode of evolutionary activity involving further growth and reorganization of the brain which resulted in an advanced theory of mind and a creative, recursive System 2 type cognitive processing and the emergence of *Homo sapiens*.

Modern, complex language emerged from the precursor protolanguage to be utilized to represent externally this mind internal universal cognition. This is the conception of language in the Representational Hypothesis (e.g. Burton-Roberts, 2011) in which there is a clear distinction between a mind internal structured cognition – the sole locus of semantics and syntax (Carey, 2011; Fodor, 1975, 2008; Harnard, 2010; Wynn et al., 2009) – and a phonological system for its semiotic representation. Despite appearances to the contrary, language is not, I claim, special in the sense of being a hardwired module of the mind. Indeed, the posited evolution of such a module presents severe strains on plausibility as evidenced in the weaknesses which Fitch, Hauser and Chomsky (2005) and Pinker and Jackendoff (2005) detect in each other's hypotheses. What *is* special, and what evolved in our species as the human faculty for language, is the

capacity for representing thoughts within the constraints of symbolic systems. Languages have the appearance of hierarchy and recursion by virtue of representing a system, of conceptual structure, that is truly hierarchical and recursive.

- Beyene, Y., Katoh, S., Woldegabriel, G., Hart, W. K., Uto, K., sudo, M., Kondo, M., Hyodo, M., Renne, P. R., suwa, G., & Asfaw, B. (2012). The characteristics and chronology of the earliest Acheulean at Konso, Ethiopia. *Proceedings of the National Academy of Science* 110, 5, 1584–1591.
- Burton-Roberts, N. (2011). On the grounding of syntax and the role of phonology in human cognition. *Lingua*, 121, 2089-2102.
- Carey, S. (2011). Precis of The Origin of Concepts. *Behavioral and Brain Sciences*. 34, 113-U117
- Eagleman, D.(2011). Incognito : the secret lives of the brain. Edinburgh: Canongate.
- Evans, J. S. B. T. (2010). *Thinking twice: two minds in one brain*. Oxford: Oxford University Press.
- Fitch, W. T., Hauser, M. D. and Chomsky, N. (2005). The evolution of the language faculty: Clarifications and implications. *Cognition*. 97, 179-210.
- Fodor, J. A. (1975). The language of thought. New York: Crowell
- Fodor, J A. (2008). LOT 2 : The language of thought revisited. Oxford: Oxford University Press
- Frankish, K. (2009). Systems and levels: Dual-system theories and the personalsubpersonal distinction. In J. S. B. T. Evans & K. Frankish. (Eds.) In two minds: dual processes and beyond. Oxford: Oxford University Press.
- Gould, S. J. & Eldredge, N. (1993). Punctuated equilibrium comes of age. Nature, 366, 223-227.
- Harnad, S. (2010). Eliminating the "concept" concept. *Behavioral and Brain Sciences*. 33, 213-+
- Jackendoff, R.(2002). *Foundations of language : brain, meaning, grammar, evolution.* Oxford: Oxford University Press.
- Kamm, G. B., Pisciottano, F. Kliger, R. & Franchini, L. F. (2013). The Developmental Brain Gene NPAS3 Contains the Largest Number of Accelerated Regulatory Sequences in the Human Genome. *Molecular Biology and Evolution*. 30, 1088-1102
- Levinson, S. C. & Evans, N. (2010). Time for a sea-change in linguistics: Response to comments on 'The Myth of Language Universals'. *Lingua*, 120, 2733-2758.
- Lynch, G. & Granger, R. (2008). *Big brain: the origins and future of human intelligence*. Basingstoke: Palgrave Macmillan.
- Penn, D. C., Holyoak, K. J. & Povinelli, D. J. (2008). Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences*, 31, 109-+.
- Pinker, S. and Jackendoff, R. (2005). The faculty of language: what's special about it? Cognition, 95 (2), 201-236.
- Tallerman, M. (2012). Protolanguage. In: M. Tallerman, M. & K. R. Gibson, (Eds.) The Oxford Handbook of Language Evolution. Oxford: Oxford University Press.

- Wynn, T. (2012). The Palaeolithic Record. In: M. Tallerman, M. & K. R. Gibson (Eds.) *The Oxford Handbook of Language Evolution*. Oxford: Oxford University Press.
- Wynn, T., Coolidge, F. L. & Bright, M. (2009). Hohlenstein-Stadel and the Evolution of Human Conceptual Thought. *Cambridge Archaeological Journal*. 19, 73-83

SEMANTIC CONDITIONING IN INTERACTION AND TRANSMISSION

OLGA FEHÉR^{*1}, ELIZABETH WONNACOTT², HANNA JÄRVINEN³, KENNY SMITH³

*Corresponding Author: o.feher@warwick.ac.uk
¹Psychology Department, University of Warwick, Coventry, United Kingdom
²Division of Psychology and Language Sciences, University College London, London, United Kingdom
³Linguistics and English Language, Edinburgh, United Kingdom

A central question in language evolution research is how fundamental properties of language have evolved and how that evolutionary process is shaped by human cognition. One property observed in all natural languages is variation. Linguistic variation tends not to be random and fully unpredictable. Rather, it is conditioned on the linguistic or social environment (Givón, 1985): linguistic or social context deterministically or probabilistically predicts the use of linguistic variants.

Previous research has shown that when children acquire artificial languages containing unpredictable variation, they often eliminate the variation by overusing one of the variants (e. g. Hudson Kam & Newport, 2009). However, at present there is no satisfying experimental account of why natural languages should contain so much conditioned variation or how conditioning comes about. We investigated the evolution of conditioned variation using an artificial language paradigm that included transmission and interaction.

We presented participants with images of objects accompanied by their descriptions in an artificial language. Depending on experimental condition, the objects were drawn from either one semantic category (e.g. all objects were animals) or two semantic categories (a mix of animals and vehicles). Each description consisted of a nonsense verb, a noun for the object and, for scenes involving multiple objects, a variable plural marker. The plural was marked by one of two markers (e.g. *dak* and *fip*) which occurred equally frequently in our initial experimenter-designed languages.

After training, participants were first asked to produce the descriptions they learnt for images (*recall phase*), and then they used the language to play a communication game with another participant or a simulated partner (*interaction phase*). We measured the use of the plural markers during recall and interaction, and in particular tracked whether variation was preserved in the miniature language, became conditioned (on lexical or semantic context) or was eliminated (by one marker overtaking the other).

In Experiment 1 we looked at the consequences of transmission for the evolution of conditioned variation. Participants were trained and tested on an artificial language as described above; we then took the recall data from one participant and used it as the training data for another participant, in a standard Iterated Learning paradigm. Conditioned variation evolved through this transmission process: in One Category chains (i.e. where stimuli were drawn from one semantic category) variation became lexically conditioned, whereas in Two Category chains, variation became conditioned on semantic category (i.e. animals took one marker, vehicles the other).

In Experiment 2 we ran transmission chains where each generation consisted of *two* participants, whose combined output was used as input for the next generation. In one condition, we used their recall data as input. Stable conditioned variation did not evolve here: in both One Category and Two Category chains the mixing of output from multiple independent individuals blocked any cumulative conditioning, replicating the findings of Smith et al. (2017). In another condition, the combined output produced by each pair during interaction (rather than recall) formed the input to the next generation. Although interaction allowed pairs to converge on a shared system, conditioned variation still did not evolve: One Category chains tended to eliminate one of the two variants, whereas Two Category chains retained both variants in free variation.

Finally, in Experiment 3 we imposed a bottleneck on transmission: participants were tested on novel nouns, forcing them to generalise their system of plural marking. Although semantically-conditioned systems of variation would allow generalisation, they did not evolve in transmission chains featuring more than one participant per generation; rather, variation was gradually eliminated.

These findings suggest that individual learners possess a bias for semantic conditioning (as observed in Experiment 1), but this bias does not straightforwardly produce stable semantically-conditioned patterns of variation; rather, systems lacking in variability are generally preferred. The conditions under which the patterns of conditioned variation we see in natural language therefore remain to be identified.

- Givón, T. (1985). Function, structure, and language acquisition. In D. Slobin (Ed.). *The crosslinguistic study of language acquisition* (pp. 1005-1028). Hillsdale, NY: Lawrence Erlbaum.
- Hudson Kam, C. L. & Newport, E. L. (2009). Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology*, *59*, 30-66.
- Smith, K., Perfors, A., Fehér, O., Samara, A., Swoboda, K., & Wonnacott, E. (2017). Language learning, language use and the evolution of linguistic variaiton. *Philosophical Transactions of the Royal Society B*, 372, 20160051.

THE COEVOLUTION OF DATA AND HYPOTHESES IN BAYESIAN CULTURAL EVOLUTION

VANESSA FERDINAND

Santa Fe Institute, Santa Fe, USA vanessa@santafe.edu

Culture is a special evolutionary system that is composed of two types of replicators: public structures in the world, such as artifacts and behaviors, and private structures in the mind, such as brain states or grammars. In this paper, I utilize a mathematical equivalence between replicator dynamics and Bayesian inference to specify a model where cultural artifacts and learners' hypotheses co-evolve.

1. A Replicator Dynamics Model of Bayesian Cultural Evolution

1.1. Model A: Artifact Evolution

The replicator dynamics equation (Hofbauer & Sigmund, 1998) is a general model of natural selection that can be used to describe the evolution of a population of cultural artifacts or behaviors over time, where $p(d_i)$ is the proportion of the *i*th artifact in the population of artifacts and $p(d_i)'$ is its proportion in the next generation. The relationship between $p(d_i)$ and $p(d_i)'$ is

$$p(d_i)' = \frac{p(d_i)f(d_i)}{\sum_{k=1}^{n} p(d_k)f(d_k)}$$
(1)

where $f(d_i)$ is the fitness of the *i*th artifact and $\sum_{k=1}^{n} p(d_k) f(d_k)$ is the mean fitness of all the artifacts in the population. $f(d_i)$ can be unpacked into the expression $f(d_i|e_j)$, which specifies that the fitness of *d* is dependent upon the environment, *e*, at any given time. The population of artifacts can adapt to a fixed environment, where $f(d_i)$ is the same each generation, or a changing environment, where $f(d_i)$ is different each generation.

1.2. Model B: Bayesian Hypotheses Evolution

Bayesian inference can be used to describe learning, where an agent entertains a set of hypotheses about the state of the world and then updates the probability of these hypotheses after observing some data from the world. The agent assigns each hypothesis a certain probability of being correct before learning, $p(h_j)$, and after learning, $p(h_j)'$. The relationship between $p(h_j)$ and $p(h_j)'$ is

$$p(h_j)' = \frac{p(h_j)p(d_i|h_j)}{\sum_{k=1}^n p(h_k)p(d_i|h_k)}$$
(2)

where $p(d_i|h_j)$ is the likelihood of the observed data under the *j*th hypothesis and $\sum_{k=1}^{n} p(h_k) p(d_i|h_j)$ is the mean probability of the data under all of the agent's hypotheses. Bayesian updating can be iterated over a sequence of observations. These observations can be the same each generation, where the hypotheses adapt to a fixed environment, or they can can differ each generation, where the hypotheses adapt to a moving target.

1.3. The Fitness Landscape of Learning

Even though Bayesian updating is a model of learning, we can talk about the evolution of hypotheses because the replicator dynamics equation and Bayesian inference are formally equivalent, as noted by (Shalizi et al., 2009; Harper, 2009).¹ The beauty of this equivalence lies in the interpretation of fitness in learning: the fitness of hypotheses (i.e. what causes them to gain differential support in a learner's mind) is simply the likelihood of the data under each hypothesis. Likewise, data are differentially reproduced on the basis of the support they have under each hypothesis: data that make more sense under a given hypothesis are more likely to survive than data points that make less sense under it. Therefore, the fitness of the hypotheses and the fitness of the data are both determined by the likelihood of the data given the hypotheses, such that:

$$f(d_i) = p(d_i|h^*) \tag{3}$$

$$f(h_j) = p(d^*|h_j) \tag{4}$$

Where the * indicates one particular hypothesis or data type. This has the effect of yoking the fitness values of the data and hypotheses to the same fitness landscape, defined by the likelihood matrix W_{ij} . Equation 5 gives an example matrix for a population of three data points and three hypotheses:

¹Due to this equivalence, we can conceptualize the probability distribution over hypotheses as a population of hypotheses. Therefore, this paper will refer to p(h) in two ways: as the probability of a hypothesis and as the proportion of a hypothesis in a population of hypotheses.

$$W_{ij} = p(d_i|h_j) = \begin{bmatrix} h_1 & h_2 & h_3 \\ .8 & .2 & .3 \\ .1 & .6 & .3 \\ .1 & .2 & .4 \end{bmatrix} \begin{array}{c} d_1 \\ d_2 \\ d_3 \end{array}$$
(5)

For example, if the population of hypotheses consists only of h_1 , the relative fitness of $\{d_1, d_2, d_3\} = \{.8, .1, .1\}$ and if the population of data consists only of d_1 , the relative fitness of $\{h_1, h_2, h_3\} = \{.62, .15, .23\}$. For any population of hypotheses that contain a mixture between hypotheses types, I define the fitness of each data type to be the weighted average of its fitness under each hypothesis, where each weight equals the hypothesis' proportion in the population. Likewise, the relative fitness of each hypothesis is the weighted average of its fitness under each data type, where weights equal each data type's proportion in the population.

1.4. Model C: Coevolution of Artifacts and Hypotheses

Model A and B can be combined to model the coevolution of data and hypotheses, where a population of artifacts adapts to a distribution over hypotheses, and vice a versa:

$$p(d_i)' = \sum_{j} \left[\left(\frac{p(d_i)p(d_i|h_j)}{\sum_{k=1}^{n} p(d_k)p(d_k|h_j)} \right) p(h_j) \right]$$
(6)

$$p(h_j)' = \sum_{i} \left[\left(\frac{p(h_j)p(d_i|h_j)}{\sum_{k=1}^{n} p(h_k)p(d_i|h_k)} \right) p(d_i) \right]$$
(7)

Equation 6 is equivalent to Equation 1, where $f(d_i)$ is specified as $p(d_i|h_j)$ and where the updated proportion $p(d_i)'$ is a sum of its updated proportion under each hypothesis, weighted by the proportion of the hypotheses at the current time step. Likewise, Equation 7 is equivalent to Equation 2, where the updated proportion $p(h_j)'$ is a sum of its updated proportion under each data type, weighted by the proportion of each data type in the current population.

2. Example Dynamics of the Model

Figure 1 provides a visualization of the model's behavior for an example set of parameters: an initial distribution over data $d_{init} = p(d_1, d_2, d_3) = \{.2, .2, .6\}$ and hypotheses $h_{init} = p(h_1, h_2, h_3) = \{.3, .4, .3\}$. The fitness values used are those in Equation 5. The top row (*Model A*) shows a population of data adapting to h_{init} , which is held fixed. The data converges to a population distribution of $\{0.56, 0.44, 0\}$, where d_1 constitutes 56% of the population, d_2 constitutes 44% of the population, and d_3 is extinct. The middle row (*Model B*) shows a population of hypotheses adapting to d_{init} , which is held fixed. The data converges to a population distribution of $\{0.56, 0.44, 0\}$, where d_1 constitutes 56% of the population, d_2 constitutes 44% of the population, and d_3 is extinct. The middle row (*Model B*) shows a population of hypotheses adapting to d_{init} , which is held fixed. The hypotheses converge



Figure 1. Example behavior of each model on a simplex plot. Axes show the proportion of each variant in the population. Corners indicate the fixation of one variant and are labeled by that variant (ex: a point in the "d1" corner means the d_1 variant comprises 100% of the population). *Model A*: Data replicate while the distribution over hypotheses remains fixed. *Model B*: Hypotheses replicate while the distribution over data remains fixed. *Model C*: Data and hypotheses coevolve, fixating in a different stable state than either variant would have on its own.



Figure 2. Example evolutionary trajectories from *Model C*, demonstrating how the length of time it takes a population to reach its stable state varies in relation to landscape typography. The y axis shows the proportion of each replicator type, which was initialized at $p(d_1, d_2, d_3) = \{.6, .2, .2\}$ and $p(h_1, h_2, h_3) = \{\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\}$. Darker colors denote peakier landscapes.

to a population composed only of h_3 . The bottom row shows the behavior of Model C, where the population of data (left) and the population of hypotheses (right) coevolve with one another. The data were updated first and therefore, the first step in the trajectory is identical to their first step in *Model A*. However, the first step the hypotheses take is not identical to their first step in *Model B* because the updated distribution over data caused the relative fitness among hypotheses to change, sending them in a slightly different direction. Although their trajectory is now altered, the hypotheses still continue toward h_3 until the population of data becomes sufficiently different to have changed the boundaries of the basin of attraction for h_1 . At this point, the hypotheses find themselves in a new basin which converges to a different attractor: $p(h_1, h_2, h_3) = \{1, 0, 0\}$.

In the replicator dynamics equation, convergence rates are an increasing function of fitness strength: as the differential growth rate among replicators becomes larger, the population reaches fixation faster. This applies to the first two models described in this paper, but not to *Model C*. Figure 2 shows six evolutionary trajectories from *Model C* which vary only in the peakiness of their fitness landscapes (where peaky entails strong differential fitness). Flatter fitness landscapes are shown in lighter grey and peakier fitness landscapes are shown in darker grey. On the left, an initial data distribution of $p(d_i) = \{.6, .2, .2\}$ converges to $p(d_i) = \{1, 0, 0\}$. On the right, an initial uniform distribution over hyptheses $p(h_j) = \{\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\}$ converges to $p(h_j) = \{1, 0, 0\}$. Trajectories on the flattest landscape have the longest convergence time, reaching fixation slower than the trajectories on slightly peakier landscapes. However, as the peakiness continues to increase, we see a reversal, with the peakiest landscapes leading back to longer and longer convergence times.



Figure 3. Example behavior of the Griffiths and Kalish (2007) model, for the likelihoods in Equation 5, a prior distribution over hypotheses $p(h_1, h_2, h_3) = \{.3, .4, .3\}$, and three different initial distributions over data. *Right*: In all cases, the posterior distribution over hypotheses converges to the prior. *Left*: In all cases, the distribution over data converges to $p(d_1, d_2, d_3) = \{.41, .36, .23\}$, which is the distribution over data that is most likely under this particular prior.

3. Comparison to Existing Models of Bayesian Cultural Evolution

In the first model of Bayesian cultural evolution, Griffiths and Kalish (2007) explored a population of Bayesian agents that shared the same, fixed prior over hypotheses. These agents receive data from one another, compute a posterior distribution, sample a hypothesis from the posterior, and then sample data from that hypothesis. Over time, the posterior distribution that agents compute converges to the prior distribution and the data converges to the integrated probability of the data under the prior distribution of hypotheses (which is the distribution over data that uniquely leads to the computed posterior being equal to the prior). Figure 3 shows an implementation of the Griffiths and Kalish model for the likelihoods in Equation 5. Three different initial distributions over data are shown. Each initial condition results in a trajectory that converges to the same point. In this model, agents produce data de novo each generation: in no sense are they copying data from the previous generation. This dynamic gradually erodes the population's information about its start state, freeing the system to converge to a unique stable state: the prior distribution over hypotheses and its corresponding data distribution. (The same behavior holds for their MAP model, with the exception that the hypotheses converge to a point near the prior.)

In another model of Bayesian cultural evolution, Beppu and Griffiths (2009) implemented a population of agents that receive data not from one another, but from a fixed source in the world. They compute a posterior distribution and then

communicate complete information about the posterior to the next generation of agents, who adopt this posterior distribution as their own prior distribution. This model is formally equivalent to a Bayesian updating model of learning in a single individual and thus, is equivalent to *Model B* (see Figure 1 for example dynamics of the Beppu and Griffiths model, where hypotheses update in response to a fixed source of data).

The main difference between these two existing models and the coevolutionary model lies in their number of unique stable states and the effect of initial conditions. The models described above each converge to one stable state regardless of their initial conditions. In the coevolutionary model, there are a minimum of n stable states representing the fixation of each variant x in the set $\{x_1, x_2, ..., x_n\}$. This holds on any given fitness landscape. For a certain class of fitness landscapes, where the fitness matrix is symmetrical about its diagonal, there is an additional stable state in the center of the simplex when both $d_i = \{\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\}$ and $h_j = \{\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\}$.

In the Griffiths and Kalish model, the evolving population gradually obtains information about the prior distribution and the speed of the information gain is determined by the likelihood structure: flatter likelihoods yield faster convergence rates (and uniform likelihoods yield immediate convergence to the prior). In the Beppu and Griffiths model, the evolving population gradually obtains information about the world (via the data received) and the speed of information gain is also determined by the likelihood structure. However, the direction of the relationship here is reversed: flatter likelihoods yield slower convergence rates (and uniform likelihoods result in no change at all). In the model where data and hypotheses co-evolve, it is less clear what the populations are gaining information about. Populations appear to be gaining information about the absorbing state of the basin of attraction they are currently in. However, as the basins of attraction change, populations switch to gain information about new absorbing states, overwriting what they previously "learned". In this model, the effects of fitness typography pattern with the Beppu and Griffiths model: flatter likelihoods yield slower convergence rates and uniform likelihoods result in no change at all.

4. Discussion and Extensions

This paper introduced a new model for exploring the coevolution of data and hypotheses in cultural evolution. I believe that Bayesian models of cultural evolution provide one possible way of formalizing Sperber (1996)'s concepts of public representations (as data) and private representations (as hypotheses). Analysis of these models could also provide quantitative insight into some of the interesting dynamics described in Cultural Attraction Theory (Sperber, 1985, 1996). Early computational models of language evolution contain many examples of private replicators (as grammars in the agents' minds) and public replicators (as utterances or strings produced from the grammar) changing in consort with one an-

other and converging to a rich diversity of semi-stable states (e.g. Kirby, 2001; Brighton, 2002). The current model was developed in order to broaden the range of evolutionary dynamics that Bayesian models of cultural evolution can capture. The next steps include fitting this model to experimental and simulation data from dynamic-rich examples of cultural evolution such as those found in social learning theory, cultural attraction theory, and the evolution of language. Various extensions to this model can be made along these lines. For example, part of the data distribution can be anchored in an unchanging world, or part of the hypotheses distribution can be anchored in an unchanging inductive bias. Furthermore, questions about causal primacy in cultural evolution could be addressed using this model. Do private replicators tend to drive the evolution of public replicators, or vice a versa? It is quite possible that certain classes of asymmetrical fitness landscapes may create dynamics where the hypotheses steer the evolution of the data and other classes of asymmetry may cause the data to steer the evolution of the hypotheses. Knowledge of this relationship could prove useful to researchers in cultural evolution who are working in domains where the likelihood relationships between public and private representations are known.

Acknowledgements

This research was supported by the Omidyar Fellowship and the Santa Fe Institute. Special thanks to Nihat Ay, Michael Lachmann, Thomas Morgan, and the postdocs at SFI for their feedback on earlier drafts of this project.

References

- Beppu, A., & Griffiths, T. (2009). Iterated learning and the cultural ratchet. In *Proceedings of the cognitive science society* (Vol. 31).
- Brighton, H. (2002). Compositional syntax from cultural transmission. *Artificial life*, 8(1), 25–54.
- Griffiths, T., & Kalish, M. (2007). Language evolution by iterated learning with Bayesian agents. *Cognitive Science*, *31*, 441–480.
- Harper, M. (2009). The replicator equation as an inference dynamic. *arXiv* preprint arXiv:0911.1763.
- Hofbauer, J., & Sigmund, K. (1998). *Evolutionary games and population dynamics*. Cambridge University Press.
- Kirby, S. (2001). Spontaneous evolution of linguistic structure-an iterated learning model of the emergence of regularity and irregularity. *Evolutionary Computation, IEEE Transactions on, 5*(2), 102–110.
- Shalizi, C. R., et al.. (2009). Dynamics of bayesian updating with dependent data and misspecified models. *Electronic Journal of Statistics*, *3*, 1039–1074.

Sperber, D. (1985). On anthropological knowledge. Cambridge University Press.

Sperber, D. (1996). *Explaining culture: A naturalistic approach* (Vol. 323). Blackwell Oxford.

THE EVOLUTION OF OPTIMIZED LANGUAGE IN THE LIGHT OF STANDARD INFORMATION THEORY

Ramon Ferrer-i-Cancho*1 and Christian Bentz^{2,3}

*Corresponding author, rferrericancho@cs.upc.edu
¹Department of Computer Science, Universitat Politècnica de Catalunya, Barcelona, Catalonia
²Department of General Linguistics, University of Tübingen, Tübingen, Germany
³DFG Center for Advanced Studies, University of Tübingen, Tübingen, Germany

Zipf's law of abbreviation, the tendency of more frequent words to be shorter, emerges as a universal property of languages (Bentz & Ferrer-i-Cancho, 2016). Language users have been shown to adhere to it as a result of combining two conditions: accuracy, i.e. avoiding ambiguity, and "efficiency", i.e. using word forms as short as possible (Kanwal, Smith, Culbertson, & Kirby, 2017). Random typing has been suggested as a non-functional alternative to statistical laws of language (Miller & Chomsky, 1963; Li, 1992). However, a caveat of random typing is that it is not purely non-functional from an information theoretic perspective: random typing can be seen as a case of optimal nonsingular encoding of information (Ferrer-i-Cancho, Bentz, & Seguin, 2015). This is an example of how information theory and evolutionary linguistics can develop mutually enriching connections.

Another example is the striking similarity between the two conditions above and coding theory, where the problem of compression is the problem of minimizing L, the mean length of codes (e.g., words) under some coding scheme (Cover & Thomas, 2006). The minimization of L matches the "efficiency" condition, whereas nonsingular coding is a coding scheme that matches beautifully the accuracy condition. Nonsingular coding is equivalent to using unambiguous words to represent meanings (Cover & Thomas, 2006, p. 105). Extensions of standard information theory predict that in case of optimal coding (maximum "efficiency" and maximum accuracy), the correlation between word frequency and word length *cannot* be positive and, in general, it is expected to be negative in concordance with Zipf's law of abbreviation (Ferrer-i-Cancho et al., 2015).

Given the alphabet of a language and the probabilities of the word types, we can calculate L_{min} , the minimum mean word length that can be achieved assuming a certain scheme (Ferrer-i-Cancho et al., 2015). Then, we can measure the degree of optimality of a language with $\eta = L_{min}/L$. η is the so-called coding efficiency (Borda, 2011), and ranges between 0 and 1, reaching 1 in case of an optimal communication system.



Figure 1. The distribution of η in the PBC under nonsingular coding (blue) and uniquely decodable coding (orange). Dashed lines correspond to a prefix of 10^5 tokens of the original text to reduce the estimation bias due to differences in text length.

Preliminary analyses of more than 1000 languages in the PBC, the Parallel Bible Corpus (Mayer & Cysouw, 2014), suggest that real languages are optimized to a 30% average (Fig. 1). Interestingly, the average optimization ratio increases to 40% (Fig. 1) if the nonsingular coding scheme is replaced by uniquely decodable encoding, which requires not only that there is no ambiguity between word types as in nonsingular coding, but also that a concatenation of letters without blanks allows for just one segmentation into a sequence of word tokens.

Such a mixture of suboptimalities in languages, which are neither perfectly nonsingular nor perfectly uniquely decodable, provides support for the hypothesis that Zipf's law for word frequencies stems from a competition between optimal nonsingular coding and optimal uniquely decodable coding (Ferrer-i-Cancho, 2016). This account is not just one more model of Zipf's law. Compression also predicts Zipf's law of abbreviation as reviewed above, as well as Menzerath's law (Gustison, Semple, Ferrer-i-Cancho, & Bergman, 2016). Hence, it illustrates the combination of predictive power, parsimony and mathematical rigor that information theory offers to understand how languages evolve universal properties.
Acknowledgements

R.F.C is funded by the grants 2014SGR 890 (MACDA) from AGAUR (Generalitat de Catalunya) and the grant TIN2017-89244-R from MINECO. C. B. is funded by the German Research Foundation (DFG FOR 2237) and by the ERC Advanced Grant 324246 EVOLAEMP.

- Bentz, C., & Ferrer-i-Cancho, R. (2016). Zipf's law of abbreviation as a language universal. In C. Bentz, G. Jger, & I. Yanovich (Eds.), *Proceedings of the Leiden Workshop on Capturing Phylogenetic Algorithms for Linguistics*. University of Tbingen.
- Borda, M. (2011). *Fundamentals in information theory and coding*. Berlin: Springer.
- Cover, T. M., & Thomas, J. A. (2006). *Elements of information theory*. New York: Wiley. (2nd edition)
- Ferrer-i-Cancho, R. (2016). Compression and the origins of Zipf's law for word frequencies. *Complexity*, 21, 409-411.
- Ferrer-i-Cancho, R., Bentz, C., & Seguin, C. (2015). Compression and the origins of Zipf's law of abbreviation. *http://arxiv.org/abs/1504.04884*.
- Gustison, M. L., Semple, S., Ferrer-i-Cancho, R., & Bergman, T. (2016). Gelada vocal sequences follow Menzerath's linguistic law. *Proceedings of the National Academy of Sciences USA*, 13, E2750-E2758.
- Kanwal, J., Smith, K., Culbertson, J., & Kirby, S. (2017). Zipfs law of abbreviation and the principle of least effort: Language users optimise a miniature lexicon for efficient communication. *Cognition*, 165, 45-52.
- Li, W. (1992). Random texts exhibit Zipf's-law-like word frequency distribution. *IEEE T. Inform. Theory*, *38*(6), 1842-1845.
- Mayer, T., & Cysouw, M. (2014). Creating a massively parallel bible corpus. In Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC-2014), Reykjavik, Iceland, May 26-31, 2014 (pp. 3158–3163).
- Miller, G. A., & Chomsky, N. (1963). Finitary models of language users. In R. D. Luce, R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 2, p. 419-491). New York: Wiley.

DO WE SEE WORD ORDER PATTERNS FROM SILENT GESTURE STUDIES IN A NEW NATURAL LANGUAGE?

MOLLY FLAHERTY^{*1}, MARIEKE SCHOUWSTRA¹, SUSAN GOLDIN-MEADOW²

^{*}Molly Flaherty: molly.flaherty@ed.ac.uk ¹Centre for Language Evolution, University of Edinburgh, Edinburgh, UK ²Department of Psychology, The University of Chicago, Chicago, USA

Typological analysis clearly shows that the world's languages are not evenly distributed among all logically possible patterns. Of the six possible orderings of Subject (S), Object (O), and Verb (V), SOV and SVO orders are vastly overrepresented in the world's languages. Studies on the emergence of word order regularities in silent gesture by hearing non-signers (e.g., Goldin-Meadow, et al., 2008; Gibson et al., 2013) overwhelmingly find evidence for SOV ordering. Based on this type of evidence, it has been proposed that SOV ordering is the most basic ordering from which all other orders emerged. However, semantic properties of the meanings to be conveyed also influence word order in silent gesture. For instance, for intensional events (in which the object is possibly non-existent or dependent on the action; e.g., 'man thinks of guitar', 'woman builds house') a cross-linguistic preference for SVO was found (Schouwstra & de Swart, 2014). Recent work finds that meaning-dependent word order patterns typical of silent gesture disappear under the influence of interaction (Christensen et al., 2016) and cultural transmission (Schouwstra et al., 2016), in favor of more consistent word order usage. However, in these studies, word order usage never becomes completely regular.

Here we investigate whether traces of the SOV/SVO pattern found in silent gesture can be observed in a new natural language: Nicaraguan Sign Language. This sign language, one of the youngest languages known to science, was born in the late 1970s with the founding of a new school for special education. Though instruction was in Spanish, students soon began to communicate with one another manually. As succeeding cohorts of students learn NSL, the language itself is changing rapidly. Though somewhat variable, NSL word order is strongly verb-final and predominantly SOV (Flaherty, 2014). However, these data are based exclusively on analysis of extensional events. If NSL word order is also influenced by semantic properties of the utterance's intended meaning, we would expect to see deviation from this SOV patterning.

Participants viewed a series of events depicting eight extensional events (i.e. woman pop bubble) and eight intensional events (i.e. woman blow bubble) involving the same object. Participants were asked to describe what they saw to a peer. Twenty-six NSL signers participated. All signers were exposed to NSL before age 7, upon school entry between the early 1980s and early 2000s.

When we analyzed SOV and SVO strings (which accounted for only 39% of strings with 1 verb and 2 arguments), we did *not* observe the pattern typical of silent gesture: SOV was dominant for both extensional and intensional events, and very few SVO strings were observed (13 total, 10 for intensional events). NSL's preference for verb-finalness (Flaherty, 2014) may not have allowed the SVO pattern to emerge. However, NSL signers tend to provide more detail than silent gesturers. As a result, many NSL strings were longer than strings observed in silent gesture. When we took into account all strings (including those with several verbs) and asked whether the Object preceded or followed target the



Figure 1. Proportion of extensional and intensional events containing 'Verb-Object' substrings.

Verb, we found more utterances with VO substrings (as opposed to OV) for intensional events than for extensional events (Fig 1). A logit mixed effects regression (with event type as fixed effect and random effects for item and signer) confirmed that strings containing VO were uncommon for extensional events (β =-2.9, SE=0.40, p < 0.001), but significantly more likely intensional for events $(\beta = 1.8,$ SE=0.45. p < 0.001). Thus, objects of intensional verbs are more likely to follow those verbs than are objects of extensionals not only in silent gesture, but also in an emerging sign language.

In this study, we find evidence for labdocumented word order preferences in an

emergent natural language: objects precede verbs for extensional events, but follow verbs for intensional events. However, this word order pattern is manifested differently in Nicaraguan Sign because it interacts with NSL's language-internal constraint for verb finalness. A combination of lab and fieldbased methodologies made this finding possible: without laboratory results, we would not have looked at a wider semantic range of events in the field; without field data, we would not have discovered the interaction between VO ordering preference and existing natural language constraints.

Acknowledgements

Supported by a Royal Society NIF (MF), a British Academy Postdoctoral Fellowship (MS), and NIH R01-DC000491 (SGM).

- Christensen, P., Fusaroli, R., Tylén, K. (2016). Environmental constraints shaping constituent order in emerging communication systems: Structural iconicity, interactive alignment and conventionalization. *Cognition*, 146, 67-80.
- Flaherty, M. (2014). The Emergence of Argument Structural Devices in Nicaraguan Sign Language (Doctoral Dissertation). Retrieved from Proquest.
- Gibson, E., Piantadosi, S. T., Brink, K., Bergen, L., Lim, E., & Saxe, R. (2013). A noisy-channel account of crosslinguistc word-order variation *Psychological Science*, 24(7), 1079-1088.
- Goldin-Meadow, S., So, W. C., Ozyurek, A., & Mylander, C. (2008). The natural order of events: How speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences USA*, 105(27), 9163-9168.
- Schouwstra, M. & de Swart, H. The semantic origins of word order. *Cognition*, 131(3), 431-436.
- Schouwstra, M., Smith, K., & Kirby, S. From natural order to convention in silent gesture. In *The Evolution of Language: Proceedings of the 11th International Conference (EVOLANG11).* Roberts, S. G., Cuskley, C., McCrohon, L., Barceló-Coblijn, L., Fehrer, O. & Verhoef, T. (eds).

A RATIONAL MODEL OF LINGUISTIC ACCOMMODATION AND ITS POTENTIAL ROLE IN LANGUAGE SIMPLIFICATION

STELLA FRANK*1 and KENNY SMITH1

^{*}Corresponding Author: stella.frank@ed.ac.uk ¹Centre for Language Evolution, University of Edinburgh, UK

Languages with large numbers of adult learners tend to be less morphosyntactically complex than languages where adult learners are rare (Wray & Grace, 2007; Lupyan & Dale, 2010; Bentz & Winter, 2013; Trudgill, 2011). This correlation between the composition of populations and linguistic complexity is often attributed to deficiencies in adult language learning. Here we investigate an additional or alternative mechanism: rational accommodation by native speakers to non-native interlocutors.

Humans have a general aptitude for reasoning about the knowledge, beliefs and motivations of other individuals, including their linguistic knowledge (e.g. Clark, 1996; Ferguson, 1981). While our interlocutors' linguistic knowledge will often be close to our own, this may not be the case in a population with many nonnative speakers. We introduce a rational model of interactions between individuals capable of reasoning about the linguistic knowledge of others, and investigate the case of a non-native speaker interacting with an native speaker who reasons about their linguistic knowledge and accommodates accordingly. Our model shows that this accommodation mechanism can lead to the non-native speaker acquiring a language variant that is less complex than the original language.

We assume a simple model in which a language consists of a distribution over linguistic variants (e.g. past tense forms). Language simplification is modelled as regularisation, whereby the most frequent variant becomes more frequent; this corresponds to, and can be measured as, entropy reduction. We model the interaction between a non-native speaker and a native speaker as interaction between two rational (Bayesian) agents. Both agents have the same initial priors and update their beliefs about the language from data in the same way, but the non-native speaker has simply seen much less data. Within an interaction, the native speaker has a parametrisable tendency to *accommodate* to the non-native speaker: instead of simply using their own language, they use the version of the language that they believe the non-native speaker may have acquired at this stage of their learning, given limited exposure. Importantly, the native speaker does not know exactly what data the non-native has seen. Instead, the native speaker models the non-



Figure 1. Distributions of final entropy of the non-native speaker's language after 100 interactions with a native speaker, for varying degrees of accommodation (indicated by colour) and different amounts of data encountered pre-interaction by the non-native speaker (indicated by the parameter |D|). These results are for a language where there are 5 variants available, whose probability distribution in the native speaker's language is (0.5, 0.2, 0.1, 0.1); this language has entropy of 1.96, indicated by the vertical dashed line. If the native speaker does not accommodate (blue lines), the non-native eventually converges to the true language. However, if the native speaker does accommodate (yellow and green lines), the non-native speaker is likely to arrive at a language that is more regular than the native speaker's; this regularization tendency is particularly pronounced when the probability of accommodation is high and the non-native speaker has seen relatively little data prior to interaction.

native speaker's linguistic knowledge by integrating over possible datasets the non-native speaker might have seen.

Representative model results for a sample language are shown in Figure 1. While learners interacting with non-accommodating speakers eventually learn the original language, non-native speakers interacting with accommodating native speakers end up learning a more regular language. This is due to the combination of the limited exposure of the non-native individual, which results in highly skewed initial distributions and some probability of not having seen low-frequency variants (Hertwig, Barron, Weber, & Erev, 2004; Hahn, 2014), in conjunction with a native speaker who is aware of and accommodates this initial bias in the non-native speaker's input, therefore providing the non-native speaker with further data which 'locks in' their biased starting point.

This model shows that accommodation by native speakers to non-native speakers during interaction can lead to language simplification, and therefore suggests how accommodation can explain the link between population makeup and linguistic complexity. The model assumes that individuals are capable of reasoning rationally about their interlocutors' linguistic knowledge, an assumption we are currently testing empirically with human learners.

Acknowledgements

This project has received funding from the European Research Council (grant agreement No. 681942).

- Bentz, C., & Winter, B. (2013). Languages with more second language learners tend to lose nominal case. *Language Dynamics and Change*, *3*(1), 1–27.
- Clark, H. H. (1996). Using language. Cambridge University Press.
- Ferguson, C. A. (1981). 'Foreigner Talk' as the name of a simplified register. International Journal of the Sociology of Language, 1981(28).
- Hahn, U. (2014, Mar). Experiential limitation in judgment and decision. *Topics in Cognitive Science*, 6(2), 229–244.
- Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004, Aug). Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, 15(8), 534–539.
- Lupyan, G., & Dale, R. (2010, Jan). Language structure is partly determined by social structure. *PLoS ONE*, *5*(1), e8559.
- Trudgill, P. (2011). Sociolinguistic typology: Social determinants of linguistic complexity. Oxford University Press.
- Wray, A., & Grace, G. W. (2007, Mar). The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua*, 117(3), 543–578.

NON-COMMUNICATIVE FUNCTIONS CAN BE EQUALLY IMPORTANT FOR STUDIES OF LANGUAGE EVOLUTION

KOJI FUJITA

fujita.koji.3x@kyoto-u.ac.jp Graduate School of Human and Environmental Studies, Kyoto University, Kyoto, Japan

Studies of animal communication play a central role in today's evolutionary linguistics. I argue here that non-communicative behaviors of animals and humans may provide equally important insights into language evolution.

Contrary to the previously held view that language is totally unique to humans, we now understand (i) that language is a complex trait consisting of several subsystems each of which may have evolved independently in other species, and (ii) that language emerged as a result of a (re)combination of these subsystems (precursors). Among these subsystems is a hierarchical compositional syntax, which still enjoys a special status as the hallmark of human language to be found nowhere else in the animal kingdom. Syntax therefore is considered to be a perplexing barrier to the natural understanding of language evolution.

Fortunately, researchers are now beginning to overcome this barrier by carrying out experiments which purport to show that a rudimentary syntax exists in nonhuman animal communication systems. Most recently, Suzuki et al. (2016, 2017) claim that the Japanese tit (*Parus minor*) may have compositional syntax, by testing these birds in ingeniously devised experimental methods and showing that they respond distinctively to novel sequences of calls in accordance with call ordering.

However, a cursory review of their study reveals that there are still some deep gaps between human syntax and the alleged compositional syntax in birds (Fujita, 2017). Most notably, their experiments show at best that bird syntax is only narrowly based on linear order, whereas it is well attested that human

language is dependent on hierarchy instead of (or in addition to) linearity. It is also unclear whether bird syntax has such typical structural properties of human syntax as recursiveness and endocentricity (headedness). Studies of primate alarm calls point to more or less a similar conclusion (Schlenker et al., 2016). In a nutshell, animal communication can be combinatorial but not compositional, as far as available evidence goes. As usual, absence of evidence is not evidence of absence, and further attempts to search for genuinely compositional syntax in nonhuman animals are highly welcome. In light of current situation, I make the following two points to make inquiries into language evolution even more integrative, constructive and fruitful.

(1) Animal communication should be better conceived of as a precursor to protolanguage with linear grammar, rather than to full human language with hierarchical grammar (Bickerton, 2014, Jackendoff & Wittenberg, 2016). This allows one to focus on the shift from linear grammar to hierarchical grammar in understanding how human language emerged. Lack of true compositionality in animal communication hints at its similarity with exocentric compounds in human language (like *birdbrain* in its metonymic sense; Jackendoff, 2009, Progovac, 2015), but as I will argue there is still a remarkable difference in that exocentric compounds in human language involves endocentricity or semantic compositionality in nature.

(2) To explain the shift/expansion from linear grammar to hierarchical grammar, one needs to look beyond animal communication and investigate non-communicative behaviors of other animals. Researchers disagree with respect to whether the original function of language was communication or not, but to say the least, given the non-monolithic nature of human language, it is unlikely that every subsystem of language evolved as an adaptation to communication from the beginning. In particular, the precursor to hierarchical grammar may be found in non-communicative functions like primitive tool use and tool making, shared by humans and nonhumans alike (Boeckx & Fujita, 2014, but contra Berwick & Chomsky, 2017).

This is plausible, for example, in light of the common neural substrates for syntax and action, and also the formal parallelism between syntactic structure building and hierarchical object manipulation. That communication may not be very relevant to the emergence of hierarchical grammar is also supported by the observation that hierarchical grammar is rather dysfunctional in terms of communicative efficiency (consider how even simple expressions like *green tea cup* or *John said Mary laughed again* can be structurally and semantically

ambiguous, a communicative disadvantage which would be absent in linear grammar).

Healthy growth of evolutionary linguistics requires a much broader perspective which pays serious attention to both communicative and noncommunicative behaviors of humans and other animals alike.

Acknowledgements

This work has been supported by MEXT/JSPS Grant-in-Aid for Scientific Research on Innovative Areas #4903 (Evolinguistics), JP17H06379, and JSPS Grant-in-Aid for Scientific Research (C) JP16K02765.

- Berwick, R. C., & Chomsky, N. (2017). Why only us: recent questions and answers. *Journal of Neurolinguistics*, 43B, 166-177.
- Bickerton, D. (2014). *More than nature needs: language, mind, and evolution*. Cambridge, MA: Harvard University Press.
- Boeckx, C., & Fujita, K. (2014). Syntax, action, comparative cognitive science, and Darwinian thinking. *Frontiers in Psychology*, *5*, 627.
- Fujita, K. (2017). On the parallel evolution of syntax and lexicon: a Merge-only view. *Journal of Neurolinguistics*, 43B, 178-192.
- Jackendoff, R. (2009). Compounding in the parallel architecture and conceptual semantics. In R. Lieber & P. Štekauer (Eds.), *The Oxford handbook of compounding* (pp. 105-128). Oxford: Oxford University Press.
- Jackendoff, R., & Wittenberg, E. (2016). Linear grammar as a possible steppingstone in the evolution of language. *Psychonomic Bulletin & Review*, 24, 219-224.
- Progovac, L. (2014). Evolutionary syntax. Oxford: Oxford University Press.
- Schlenker, P., Chemla, E., & Zuberbühler, K. (2016). What do monkey calls mean? *Trends in Cognitive Sciences*, 20, 894-904.
- Suzuki, T. N., Wheatcroft, D., & Griesser, M. (2016). Experimental evidence for compositional syntax in bird calls. *Nature Communications*, 7,10986.
- Suzuki, T. N., Wheatcroft, D., & Griesser, M. (2017). Wild birds use an ordering rule to decode novel call sequences. *Current Biology*, 27, 2331-2336.

PRIMATE SONGS AND THEIR RELEVANCE IN THE STUDY OF LANGUAGE EVOLUTION

MARCO GAMBA^{*1}, VALERIA TORTI¹, LONGONDRAZA MIARETSOA¹, DARIA VALENTE¹, CHIARA DE GREGORIO¹, MARTINA TUBITO², MARTINA ZARANTONELLO¹, VITTORIO L. BIANCO¹, LIVIO FAVARO¹, OLIVIER FRIARD¹, and CRISTINA GIACOMA¹,

*Corresponding Author: marco.gamba@unito.it ¹Department of Life Sciences and Systems Biology, University of Torino, Torino, Italy ²Zoom Torino Spa, Cumiana (To), Italy

It has been suggested that early human vocal communication should result from at least two critical abilities: the ability to engage in cooperative efforts (Levinson, 2005), and the ability to combine sounds in larger structures (Collier et al., 2014). These characteristics are essential features of choral songs, and, in fact, Darwin speculated that language might have originated from singing. The rhythms and pitch contours that distinguish songs in other species may be investigated to understand the roots of our language. Many authors have indicated birds and singing primates as the natural target of these studies (Geissmann, 2000).

The evolution of songs in nonhuman primates has been associated with a monogamous mating system and to the active defense of a territory. We have investigated the layers of complexity of the songs emitted by three nonhuman primate species: the siamang (*Symphalangus syndactylus*), the white-handed gibbon (*Hylobates lar*), and the indri (*Indri indri*). Although the single species have received attention from scientists, a comparative approach investigating frequency variation and individual temporal patterns has never been used.

For each of these species, we analyzed songs from different social groups and hierarchically classified the units emitted in the songs in phrases using dynamic time warping and cluster analysis. For each unit, we extracted a pitch

contour and labeled it by the context in which songs were emitted and by the sex of the emitter.

We found that a) primate songs show a species-specific structure in the organization of units and phrases; b) a shared feature across the three species was the presence of an apparent sex dimorphism, and c) song given in different context (e.g., cohesion vs. advertisement) have distinctive acoustic structures. Moreover, we found that the structure of phrases possessed individually distinctive characteristics, thus the potential to provide conspecifics with emitter's identity cues.

These findings are relevant to the evolution of language for three reasons. Similarly to humans, siamangs, white-handed gibbons, and indris assemble simple units into more complex structures to convey different information, and individuals react differently to different songs, confirming the presence of functionally referential communication systems and a simple phonological syntax in nonhuman primates (Marler, 1977). The songs reach their most consistent portions in ascending or descending sequences of units, which indeed represent an interesting case of timing and pitch variation, a crucial feature of birdsong and human speech (Levinson and Holler, 2014). The presence of pitch sex dimorphism in nonhuman primate vocal signals is rare, and it has been indicated as a critical element in the evolution of human perceptual abilities (Patel, 2010). Finally, portions of the song have the potential to provide information about the individual identity of the emitters, suggesting that the transmission of identity-related cues at long distance is a trait of the living singing primates may have in common with the first human societies (Brown, 2000).

The range of species capable of cooperative singing is very restricted, and further research is needed to find out whether the singing capacity is associated with other critical features of human communication and musicality that have rarely been found in nonhuman species (e.g., the capacity to synchronize with a beat). These future studies will contribute to the identification of which aspects of our nonverbal auditory processing are shared with other species, and which are indeed uniquely human.

References

Brown, S. (2000). The "musilanguage" model of music evolution. In N.L. Wallin, B. Merker & S. Brown (Eds.). *The Origins of Music* (pp. 271–300). Cambridge: MIT Press.

- Collier, K., Bickel, B., van Schaik, C. P., Manser, M. B., & Townsend, S. W. (2014). Language evolution: syntax before phonology? *Proceedings of the Royal Society B: Biological Sciences*, 281, 20140263.
- Geissmann, T. (2000). Gibbon songs and human music from an evolutionary perspective. In N.L. Wallin, B. Merker & S. Brown (Eds.). *The Origins of Music* (pp. 103–123). Cambridge: MIT Press.
- Levinson, S. C. (1995). Interactional biases in human thinking. In E. N. Goody (Ed.), Social Intelligence and Interaction (pp. 221-260). Cambridge: Cambridge University Press.
- Levinson, S. C., and Holler, J. (2014). The origin of human multi-modal communication. *Philosophical Transactions Royal Society London B: Biological Sciences*, 369, 20130302.
- Marler P (1977) The structure of animal communication sounds. In: Bullock TH (Ed.). *Recognition of complex acoustic signals* (pp. 17–35). Berlin: Springer.
- Patel, A. D. (2010). *Music, language, and the brain*. Oxford: Oxford university press.

HOW THE EVOLUTION OF TEACHING MAY HELP UNDERSTANDING THE EVOLUTION OF LANGUAGE

PETER GÄRDENFORS *1,2

*Corresponding Author: Peter.Gardenfors@lucs.lu.se
¹Department of Philosophy, Lund University, Lund, Sweden
²Stellenbosch Institute for Advanced Studies, Wallenberg Research Centre at Stellenbosch University, Marais Street, Stellenbosch 7600, South Africa

1. Introduction

Teaching is present in all extant human societies. On the other hand, intentional teaching within other species is very limited. In an analysis of the evolution of teaching, demonstration and pantomime come out as pivotal capacities. The aim of this paper is to spell out the relevance of this for the evolution of language.

2. The Evolution of Teaching

Gärdenfors and Högberg (2017) distinguish six levels of intentional teaching: (1) intentional approval/disapproval, (2) drawing attention, (3) demonstrating and pantomime, (4) communicating concepts, (5) explaining concept relations and (6) narrating. They hypothesize that level after level has been added during the evolution of teaching. They analyse communicative requirements for the levels, concluding that displaced communication is required for level 4 and symbolic language only for levels 5-6.

3. The Central Role of Demonstration and Pantomime

It is only humans who teach according to levels (3)–(6). A central question is therefore why only hominins have the capacity to demonstrate and pantomime, that is, to reach level (3). This capacity seems to have generated a breakthrough in hominin teaching and in transmission of culture. When a teacher demonstrates to a learner how to perform a certain task, the demonstrator actually performs the actions involved in the task, commonly using material culture. In pantomime the mimer (teacher) only performs the movements of the actions in the task. I argue that the capacity to pantomime has been pivotal in the evolution of communication.

For pantomime (but not for demonstration), the addressee must understand that the teacher intends the pantomime to *stand for* a real action and that the teacher intends the addressee to realize this. Zlatev et al. (2005) call this the

'communicative sign function'. Another important difference with respect to demonstration is that pantomime is *displaced* (Zywiczynski et al. 2016, sect. 3.8). However, pantomime is neither conventional nor symbolic.

4. From Pantomime to Language

Following Gärdenfors and Högberg (2017), I instead view the evolution of teaching as a series of stages that I argue can be matched to different stages in the evolution of communication. There are two main functions for pantomime: The primary is an invitation to copy – the teaching function. The secondary is a communicative function. This distinction also shows up when determining what is the *intention* of a pantomime. A teacher can pantomime an *action* that the teacher wants the student to *copy*, and an action can be pantomimed as part of a *message* (request, command, warning, narrative, etc.). In the gesture literature, the communicative use of pantomime has been in focus.

Pantomime has been argued to be a precursor to protosign and protolanguage. Arbib (2012) suggests that protosign develops by conventionalization out of pantomime and other gestures. A crucial feature is that pantomime provides an *open* set of gestures that can be generated to create new meanings. It should be noted that conventions presume *common knowledge* among the signers), something that requires higher order mindreading. Arbib (2012, p. 231) then proposes that protospeech builds on protosign in an 'expanding spiral'.

Gärdenfors (2013) argues that symbolic language is necessary for the advanced forms of *cooperation* that have evolved along the hominin line. Teaching should, however, also be seen as a form of cooperation and the later stages in the model of Gärdenfors and Högberg (2017) require advanced forms of communication. I conclude that pantomime is necessary for the emergence of any of these forms.

- Arbib, M. (2012). *How the Brain Got Language: The Mirror System Hypothesis*. Oxford: Oxford University Press.
- Gärdenfors, P. (2013). The evolution of semantics: Sharing conceptual domains. In *The Evolutionary Emergence of Language*, eds. Botha, R. and Everaert, M. (Oxford: Oxford University Press), 139-159.
- Gärdenfors, P. & Högberg, A. (2017). The archaeology of teaching and the evolution of *Homo docens*. *Current Anthropology*, *58*, 188-201.
- Zlatev, J., Persson, T. & G\u00e4rdenfors, P. (2005a). Bodily mimesis as 'the missing link' in human cognitive evolution. Lund: Lund University Cognitive Studies 121.
- Zywiczynski, P., Wacewicz, S. & Sibierska, M. (2016). Pantomime for language evolution research. *Topoi*, DOI 10.1007/s11245-016-9425-9.

PATTERN SIMILARITY IN BIOLOGIAL, LINGUISTIC, AND SOCIOCULTURAL EVOLUTION

NATHALIE GONTIER*1

*Corresponding Author: <u>nlgontier@fc.ul.pt</u>
¹Applied Evolutionary Epistemology Lab, Centre for Philosophy of Science, Faculty of Science, University of Lisbon, Portugal

Linguistic and sociocultural evolution models share pattern similarity with biological evolution models. Examples of similar patterns include descent with modification, reticulation, drift, gradualism, punctuated equilibria, cyclicity, and periodicity. Pattern similarity is often explained by assuming that the same evolutionary mechanisms are causally responsible for the patterns. We demonstrate that this argument is not always warranted because similar patterns can be induced by different mechanisms and processes. We investigate the implications this finding has on how we define mechanisms and on how biological, sociocultural and linguistic evolution relate and diverge from one another.

1. Introduction

Scholars are currently applying a similar set of macro-oriented phylogenetic methods (Pagel, 1999) and micro-oriented experimental evolution techniques (Mesoudi, 2016; Tamariz & Kirby 2016) or computer-generated simulations to model and mimic biological, linguistic and sociocultural evolution (Steels, 2015). Results of research indicate that biological, linguistic and cultural evolution share similar as well as diverging patterns.

Here, we focus on pattern similarities and examine the explanations given for their occurrence. Pattern similarity is found in how traits are distributed (through vertical descent with modification, horizontally and reticulately, or randomly through drift), and at what rate such distributions occur (gradually, by means of punctuated equilibria, cyclic, or periodically).

In what follows, we *first* define patterns as intermediary steps in evolutionary research that hold the middle between raw data and theoretical frameworks. Theoretical frameworks, in turn, refer to mechanisms to explain the patterns retrieved from modeling. *Secondly*, we demonstrate that pattern similarity is often explained from within the same theoretical frameworks and by referring to the same or similar evolutionary mechanisms. We investigate the validity of this assumption. *Finally*, we demonstrate that this analysis is worthwhile because it

informs us on how linguistic and sociocultural evolution relate to overall biological evolution and on how we can identify multiple units, levels, mechanisms and processes of evolution.

2. Defining Patterns

Patterns emerge from how collected data is joined into descriptive or quantitative models of events. Scholars use methodological toolkits to generate the models, and the models bring forth observable patterns that in turn are explained by referring to theoretical frameworks. Patterns are thus intermediary steps in evolutionary research that hold the middle between raw data on the one hand and explanatory theoretical frameworks on the other (Gontier 2016; Grande & Rieppel, 1994).

2.1. Patterns are Not Raw Data

Biological, linguistic and sociocultural evolution share similar patterns in (1) the *directionality of trait distribution* across lineages in time and/or space which can be vertical, horizontal, random, bi-or multi-directional, or cyclical; and (2) the *rate whereby lineages evolve* which can be gradual, fast, punctuated or periodic.

Points (1) and (2) provide a list of logical possibilities whereby we can respectively study how traits are distributed (the mode), or how we can conceptualize the rate whereby evolution occurs (the tempo), both at a micro- or macroscale (Simpson, 1944). Yet we are unfamiliar with thinking about patterns in this way. Instead, we have obtained our knowledge on the patterns of evolution through a historical learning process resulting from observing natural processes and testing hypotheses and theories on how evolution possibly occurs, as well as by experimenting how evolution can be adequately modeled (in scales of nature, timelines, trees, networks, cycles, etc.).

Patterns are informed by theories that make assumptions on how raw data should be ordered and represented. In addition, modeling by trees or networks nowadays often associates with a choice between different software packages that work from within different premises and heuristics.

Most evolutionary theories have so far focusses on how traits are distributed or transmitted across time and space, and at what rate such distribution has occurred, and that is the very reason why scholars have taken on the endeavor to model exactly these two aspects of the evolutionary process. We see this point more clearly by asking what other means there are to examine and model evolutionary phenomena.

One answer has been given by the ecological sciences that, rather than examining distribution (rates), examine how entities *interact* with each other and the abiotic world (Fox et al., 2001; Futuyma, 2010). Another answer has been given by evolutionary developmental schools that focus on how traits develop within individuals and groups (Hallgrimson & Hall, 2011). Instead of studying

phylogenetic splits, mergers or extinction, within evo-devo the emphasis lies on ontogenetic, intra-lineage phenomena. Still other means by which we can analyze evolution include hierarchy theory (Salthe, 1985), multilevel selection theory (Lewontin, 1970; Okasha, 2005) or major transitions (Maynard Smith & Szathmáry, 1995).

2.2. Patterns are Deduced from Theory-Informed Models and Diagrams

Within macro-oriented phylogenetics, the mode and tempo of evolutionary phenomena are modeled either in rooted or unrooted and bi- or multifurcating trees or networks. *Trees* are often set in a Cartesian two-dimensional coordinate system where trait distribution is tracked linearly in space and over time. Tree models, therefore, work from within the premise of vertical, often hereditary transmission and they require time consistency. For those reasons, they cannot adequately depict instances of oblique, reversed, horizontal or multidirectional transmission (Feldman & Cavalli-Sforza, 1976; Zhaxybayeva & Doolittle, 2011).

These latter phenomena are modeled in networks. *Networks* are sometimes set in vector space, and they often remain unrooted because time-consistency is not always prioritized (Morrison, 2016). Nonetheless, the length and directionality of the branches sometimes indicate time as measured by the evolutionary distance between the modelled entities. Networks model all sorts of relationships occurring within and between biological, linguistic or cultural entities. And they do so not only in the past but also in the present, thereby assuming that ontogeny and ecology have relevant evolutionary roles to play.

Interactions and developmental processes are also represented by unrooted networks, circles or overlapping Venn diagrams that track cycles (Bechtel, 2011). *Cycles* are little-appreciated patterns. They can be found in circadian rhythms, gene-regulatory networks, protein folding, overall anatomical, sexual, cognitive, linguistic, and cultural development, and aging. Cycles differ from trees and networks because the latter two model *events* (speciation, extinction, merging or splitting of branches) while cycles model *recurring* and *recursive processes*. Cycles, moreover, often show stability over time, and they sometimes follow *periodicities* (either occurring at intervals or reoccurring at specific moments during ontogeny), but no diagrams exist yet to model these periodicities.

3. Pattern Similarity and Evolutionary Frameworks

Patterns are observed in diagrams and evolutionary models that are informed by theory. Theories underlie evolutionary frameworks, and the latter, in turn, refer to mechanisms to explain the patterns of evolution. Historically, the identification of evolutionary patterns associated with the introduction of different evolutionary theories on the mode and tempo of evolution, and with an investigation into which diagrams and models are best suited to order the data deemed relevant. For those reasons, similar patterns are explained by referring to the same evolutionary, often biological mechanisms. In this part, we investigate which patterns are associated with which theories and mechanisms.

3.1. Pattern Similarity in Trait Distributions and Associated Mechanisms

In the biological sciences, the study of *vertical* trait distribution associates with selection theory that demonstrates that gene distribution occurs via a *pattern of descent with modification* driven by heredity and positive selection of adaptive (genetic) traits (Dawkins, 1983). The preferred diagrams to model this pattern are trees and cladograms. Sociocultural and linguistic sciences have also found this pattern in how linguistic and cultural traits are transmitted across generations through (un)directed or biased learning and teaching. Consequently, the pattern similarity is explained by universal selection theories that are based upon the three Darwinian principles (Lewontin, 1970): differential variation, reproduction, and selection – that can also be understood as a recurring cycle (Cavalli-Sforza & Feldman, 1981; Lumsden & Wilson, 1981; Mesoudi, 2016).

However, vertical descent with modification can also result from *random trait distributions*, otherwise known as *genetic drift*. The founders of the Modern Synthesis recognized drift to alternate with natural selection, and especially Kimura (1960) explained how drift is a mechanism that occurs independently from natural selection. However, drift theory first emerged in linguistics (Greenberg, 1960; Jesperson, 1909; Koerper & Stickel, 1980; Wittmann, 1969), and here, drift was often interpreted either as directed or as circular, and *cyclic* (Van Gelderen, 2013). Modern applications of drift theory are given by Bentley et al. (2004), Chang (2013), Centola et al. (2007), Chiaronia et al. (2009), and Koerper & Stickel (1980).

Cycles associate with both macro-oriented disciplines such as ecology (e.g. the nitrogen cycle, that also relies on symbiosis) and micro-evolutionary research fields such as evolutionary developmental biology where scholars study intra- and intergenerational developmental processes. Cycles are also studied by linguistic and sociocultural scientists. Example include the cycle whereby children learn the language they speak from their community, while they also influence the language of the community and how it is taught to them (Mesoudi, 2016); or the cycle where biological evolution underlies the development of individual and group cognition that in turn influences cultural transmission whereafter the cycle repeats because cognition and cultural transmission can influence the course of biological evolution (Maynard Smith & Szathmàry, 1995). Other examples include the iterated learning experiments and computational models that implement Bayesian statistics and Markov chain Monte Carlo simulations that mimic learning within and across generations (Briscoe, 1998; Tamariz & Kirby, 2016). The outcome of such learning is depicted in tree models, but the learning process itself is cyclic and periodic, occurring at specific moments during development over multiple generations. Cycles currently remain unintegrated into

phylogenetic representations, but for promising work on iterated learning networks, see Sole et al. (2010).

In biology, *horizontal* (from one lineage to another), *bi*- (between 2 lineages) and *multi-directional* (between multiple lineages) trait distribution has been associated with reticulate evolution as it occurs by means of lateral gene transfer, symbiogenesis, and hybridization (Arnold, 1997; Gontier, 2015; Keeling & Palmer, 2008; Margulis, 1998; Zhaxybayeva & Doolittle, 2011). In linguistic and sociocultural evolution, reticulations result from language and cultural borrowing, mixing, or hybridization (Croft, 2000; Shijulal et al., 2010; List, 2013). All these phenomena result in *reticulate patterns* which are better modeled by *networks* than trees. While especially the biological sciences have long downgraded the relevance of horizontal exchange, diachronic linguists and also anthropologists have always recognized the importance of reticulations in cultural diffusion studies.

3.2. Pattern Similarity in Distribution Rates and Associated Mechanisms

Natural selection theory traditionally predicts evolution to occur *gradually*. But the necessity of gradualness has been called into question by many scholars, including scholars that study punctuated equilibria (Gould & Eldredge, 1977).

The *pattern of punctuated equilibria* recognizes two additional patterns that often occur sequentially in time. Long periods of *stasis* are intermitted by *rapid periods* of morphological change and/or speciation. The pattern can be detected at a molecular level (Pagel et al., 2006); as well as in the archeological record (Eldredge & Tattersall, 1982), and in certain language families (Atkinson et al, 2008; Gray & Jordan, 2000; Gray et al, 2009). But although the pattern can be found to occur repeatedly within all these different phenomena, it remains an open question whether punctuated equilibria also follow *cycles* or *periodicities*. If they do, then scholars should be able to uniformly clock and predict speciation/divergence and extinction/death rates, or how long recurring periods of stasis last. This, however, remains difficult regardless the fact that stasis, rapid speciation, and extinction occur repeatedly. It might also be unwarranted to assume that any uniformity can be found in how stasis or rapid change interchange one another. Until scholars can answer the current enigmas, punctuated equilibria are, therefore, best treated as a single pattern.

The rate of *drift* is difficult to calculate and differential depending upon the traits one examines to be drifting (genes or aspects of languages and cultures) (Cavalli-Sforza & Feldman, 2003; Hallatschek et al., 2007; Nei et al., 1975). Drift can associate both with patterns of gradual descent with modification, or with punctuated equilibria. Over *short* time spans, drift often associates with a *pattern similar to Brownian motion*. Over *longer* periods of time, it either brings forth *gradual patterns* of change or *patterns of stasis*, i.e. periods wherein no reticulations or bifurcations of the lineages occur. This does not exclude the possibility for the lineage to change inside out through, for example, bottleneck

or founder effects (that correlate with environmental circumstances and population size). When drift is followed by incremental growth, it brings forth more *rapid*, *stage-like patterns*.

Reticulate evolution often occurs at a *fast* evolutionary rate, because gene, trait or organismal (e.g. microbial) transfer, as well as sociocultural and linguistic transfer can occur a-sexually (without the requirement of reproducing a next generation). Nonetheless, in nature, obligate symbiotic partnerships can also constrain evolution and either result in *co-evolution or stasis*. And in culture, individual cognition or cultural learning and teaching also provides constraints and co-evolutionary dynamics of how languages and cultures evolve.

The same goes for the *cycles* studied by developmental evolutionary biologists, psychologists, and epigenetics (Hallgrimson & Hall, 2011). Genes determine and constrain development thereby attributing to stability and *stasis*, but epigenetic changes can *rapidly* alter the course of ontogeny and phylogeny. Cycles are furthermore *periodic*, because they repeat over generations through time, and often at specific moments in time.

4. Implications for How We Understand Linguistic and Cultural Evolution

Similar patterns are often explained by the same specific biological mechanisms that subsequently become "universalized". But two problems arise. *For one,* assuming congruency between patterns and mechanisms is not always warranted because, as we saw in the previous part, the same patterns can sometimes be generated by different mechanisms. It follows that although tree and network models do lend insight into patterns of evolution, they do not straightforwardly demonstrate how, by which mechanisms, linguistic and cultural traits, organisms or species evolve. *Secondly*, when the same evolutionary theories are invoked to explain similar patterns, we find that the "universal" mechanisms held responsible become defined differentially within the different domains. Mechanism-explanations are often abandoned in favor of process accounts.

4.1. Different Mechanisms Can Induce Similar Patterns

While the pattern of descent with modification has spurred generations of researchers to investigate how bifurcation and extinction occur by means of natural selection, today we know that all mechanisms described above can bring forth this pattern. Descent with modification thus provides a general guideline to understand changing phenomena as undergoing evolution, but more research is required to examine how exactly this change occurs. Drift can, under certain circumstances, bring forth the pattern, but it can also bring forth patterns of punctuated equilibria. Reticulate evolutionary mechanisms bring forth reticulate patterns in a first phase, but in a second phase the reticulately acquired traits can undergo vertical descent with modification or bring forth patterns of co-evolution or stasis.

The take-home messages of these findings are that (1) different mechanisms can bring forth the same patterns which implies that patterns are not clairvoyant identifiers of mechanisms whereby phenomena evolve; (2) no isomorphism between patterns and specific phenomena can be adhered to.

This makes it necessary to embrace a more pluralistic account on both the nature of patterns and mechanisms. Indeed, tracking the evolution of specific lineages over long periods of time often brings forth a sequential (perhaps sometimes cyclic) series of patterns, and a single phenomenon often evolves by multiple mechanisms.

Taking our hominin past as exemplar, we now have firm evidence that, in addition to evolving by means of natural selection, our species acquired genes through hybridization with other hominin species, and through lateral gene transfer with viruses and microorganisms. Rapid reticulate patterns are intermitted with gradual descent with modification, where mutated and acquired genes are transmitted vertically. In addition, our cultural and linguistic traits evolved through a combination of guided selection, random drift, and opportune reticulation.

4.2. From Mechanism to Process Accounts

The linguistic and sociocultural sciences have taken on an evolutionary outlook by looking for patterns and mechanisms in linguistic and cultural evolution that are similar to recognized patterns and mechanisms in the biological sciences. For those reasons, scholars have reformulated gene-based natural selection theory into universal Darwinian accounts. This has introduced a shift from investigating the *mode* and *tempo* of Darwin's mechanism of natural selection to identifying a variety of *processes* that are selective.

The same can be said about reticulations. Reticulate evolution does not always require lateral gene transfer or hybridization. Reticulations also characterize processes of language and cultural borrowing or the formation of multicultural societies. This again brings forth a *shift* from mechanism to process accounts, and the similarity is found in patterns rather than in mechanisms that are specific to genetic, physiological, developmental, linguistic or sociocultural evolution.

Scholars remain divided on how sociocultural and linguistic evolution relate to biological evolution. The analysis provided here shows that a universal evolutionary approach is possible if we recognize that different *processes* can *causally* lead to pattern similarity in life, language, and culture, even without fixed mechanisms. This implies that we need to go beyond theorizing on the nature of specific mechanisms and instead focus on the wide variety of processes whereby evolution occurs because these are differential.

5. Toward Unit, Level, and Mechanism Plurality

Finally, the shift from mechanism to process accounts also alters how we define the units and levels of evolution, or more generally, how we define information. What counts as "information" in the biological, sociocultural and linguistic sciences and how "information transfer" is conceptualized is defined differentially by the different domains. These differences can also be taken as point of analysis to distinguish amongst research schools.

The gene is classically assumed to be the unit of information as well as the unit of selection. Scholars that want to reduce culture and language to biology have theorized that linguistic and sociocultural evolution can be reduced to the study of genes. Those that want to differentiate language and culture from biological evolution have often pointed out that more than genes are exchanged in linguistic and cultural evolution. And those that want to understand language, culture and biology as co-evolving realms have on the one hand searched for cultural replicators such as memes that perform functions similar to genes (Dawkins, 1983), and on the other, they have searched for other units of evolution, that surpass the scope of replicators, such as interactors (Hull, 2001) and reproducers (Griesemer, 2000).

These debates have mostly been held at a theoretical level. They do not stroke well with actual scientific practices where scholars study the transfer of a wide variety of phenomena that are understood, not only to carry information, but also to causally inflict change.

The information that linguists track extends classic etymological research and involves linguistic traits such as cognates, loan words and doublets, or linguistic universals including syntax typology (e.g. the transitions from SOV to SVO). Sociocultural scholars draw their phylogenies and networks of sociocultural evolution by tracking material artifacts ranging from paleolithic stone tools to modern skate board decks (Prentiss et al., 2016), or immaterial mentifacts such as fairy tales (da Silva & Tehrani, 2016) or religious ideas. And while biologists used to focus on the differential distribution and transmission of morphological traits amongst organisms, species and higher taxa over geological time, today, they focus more on the divergence of genes and proteins over time, which they calculate by making use of (relaxed) molecular clock models.

All domains are moving toward a general recognition that evolution, be it biological, social or linguistic in kind, occurs through a myriad of units at various levels of an ontological hierarchy, and by numerous processes rather than by a fixed set of theorized mechanisms. It are the interactions amongst the units and levels that induce evolutionary change and that bring forth pattern similarity.

Acknowledgements

Written with the support of the Portuguese Fund for Science and Technology, grant ID SFRH/BPD/89195/2012 and project number UID/FIL/00678/2013.

- Arnold, M. L. (1997). Natural hybridization and evolution. New York, NY: Oxford University Press.
- Atkinson, Q. D., Meade, A., Venditti, C., Greenhill, S. & Pagel, M. (2008). Languages evolve in punctuational bursts. *Science*, *319*, 588.
- Bechtel, W. (2011). Mechanism and biological explanation. *Philosophy of Science*, 78, 533-577.
- Bentley, R., Hahn, M. & Shennan, S. (2004). Random drift and culture change. *Proceedings of the Royal Society B*, 271, 1443–1450.
- Briscoe, E. J. (1998). Language as a complex adaptive system: Coevolution of language and of the language acquisition device. In: van Halteren, H., et al. (Eds.), *Proceedings of eighth computational linguistics in the Netherlands Conference* (pp. 3-40). The Netherlands, Amsterdam.
- Cavalli-Sforza, L. L. & Feldman, M. W. (1981). Cultural transmission and evolution. Princeton, NJ: Princeton University Press.
- Cavalli-Sforza, L. L., Feldman, M. W. (2003). The application of molecular genetic approaches to the study of human evolution. *Nature Genetics*, 33, 266-75.
- Centola, D., González-Avella, J., Eguíluz, V. & San Miguel, M. (2007). Homophily, cultural drift, and the co-evolution of cultural groups. *Journal of Conflict Resolution*, 51, 905-929.
- Chang, C. (2013). A novelty effect in phonetic drift of the native language. *Journal of Phonetics*, 41, 520-33.
- Chiaronia, J., Underhillb, P. & Cavalli-Sforza, L. (2009). Y chromosome diversity, human expansion, drift, and cultural evolution. *Proceedings of the National Academy of Science, USA, 106*, 20174-20179.
- Croft, W. (2000). *Explaining language change: An evolutionary approach*. Essex, Pearson.
- da Silva, S. G., Tehrani, J. J. (2016). Comparative phylogenetic analyses uncover the ancient roots of Indo-European folktales. *Royal Society Open Science*, 3, DOI: 10.1098/rsos.150645.
- Dawkins, R. (1983). Universal Darwinism. In Hull, D.L. & Ruse, M. (Eds.), *The philosophy of biology* (pp. 15-35). New York, NY: Oxford University Press.
- Eldredge, N. & Tattersall, I. (1982). *The myths of human evolution*. New York, NY, Columbia University Press.
- Feldman, M. & Cavalli-Sforza, L. (1976). Cultural and biological evolutionary processes, selection for a trait under complex transmission. *Theoretical Population Biology*, 9, 238-59.
- Fox, C. W., Roff, D. A. & Fairbairn, D. J. (2001). *Evolutionary Ecology: Concepts and case studies*. Oxford, Oxford University Press.
- Futuyma, D. J. (2010). Evolutionary constraint and ecological consequences. *Evolution*, 64, 1865–1884.

- Gontier, N. (2016). Guest-Editorial Introduction: Converging Evolutionary Patterns in Life and Culture. *Evolutionary Biology*, *43* (4), 427-445.
- Gontier, N. (Ed.) (2015). Reticulate Evolution. Dordrecht, Springer.
- Gould, S. J., Eldredge, N. (1977). Punctuated equilibria: The tempo and mode of evolution reconsidered. *Paleobiology*, *3*, 115–151.
- Grande, L. & Rieppel, O., (Eds.). (1994). *Interpreting the hierarchy of nature*. San Diego, CA, Academia Press.
- Gray, R. D. & Jordan, F. M. (2000). Language trees support the express-train sequence of Austronesian expansion. *Nature*, 405, 1052-1055.
- Gray, R. D., Drummond, A. J. & Greenhill, S. J. (2009). Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science*, 323, 479-483.
- Greenberg, J. H. (1960). A quantitative approach to the morphological typology of language. *International Journal of American Linguistics*, 26, 178-94.
- Griesemer, J. (2000). The units of evolutionary transition. Selection, 1, 67-80.
- Hallatschek, O., Hersen, P., Ramanathan, S., Nelson, D. R. (2007). Genetic drift at expanding frontiers promotes gene segregation. *Proceedings of the National Academy of Sciences, USA, 104 (50),* 19926-19930.
- Hallgrimsson, B. & Hall, B. (Eds.) (2011). *Epigenetics*. Berkeley, CA: University of California Press.
- Hull, D. L. (2001). *Science and Selection*. New York, Cambridge University Press.
- Jespersen, O. (1909). A modern English grammar on historical principles. London: Allen & Unwin.
- Keeling, P. J. & Palmer, J. D. (2008). Horizontal gene transfer in eukaryotic evolution. *Nature Reviews*, *Genetics*, 9, 605-618.
- Kimura, M. (1968). Evolutionary rate at the molecular level. *Nature*, 217, 624– 626.
- Koerper, H., Stickel, E. (1980). Cultural drift: A primary process of culture change. *Journal of Anthropological Research*, *36*, 463-469.
- Krause, J., et al. (2007). The derived *FOXP2* variant of modern humans was shared with Neanderthals. *Current Biology*, 17, 1908-12.
- Lewontin, R. (1970). The levels of selection. Annual Review of Ecological Systems, 1, 1-18.
- List, J.-M., et al. (2013). Networks of lexical borrowing and lateral gene transfer in language and genome evolution. *Bioessays*, *36*, 141-150.
- Lumsden, C. & Wilson, E. (1981). *Genes, mind and culture: The coevolutionary process.* Cambridge, MA: Harvard University Press.
- Margulis, L. (1998). The symbiotic planet. London: Phoenix Orion Books.
- Maynard Smith, J. & Szathmáry, E. (1995). *The major transitions in evolution*. New York, NY: Oxford University Press.
- Mesoudi, A. (2016). Cultural evolution: A review of theory, findings and controversies. *Evolutionary biology*, 43 (4), 481-497.

- Morrison, D. (2016). Genealogies: Pedigrees and phylogenies are reticulating networks, not just diverging trees. *Evolutionary Biology*, *43*, 456-473.
- Nei, M., Maruyama, T., Chakraborty, R. (1975). The bottleneck effect and genetic variability in populations. *Evolution*, 29, 1-10.
- Okasha, S. (2005). Multilevel selection and the major transitions in evolution. *Philosophy of Science*, 72, 1013-1025.
- Pagel, M. (1999). Inferring the historical patterns of biological evolution. *Science*, 401, 877-884.
- Pagel, M., Venditti, C., Meade, A. (2006). Large punctuational contribution of speciation to evolutionary divergence at the molecular level. *Science*, 314, 119-121.
- Prentiss, A. M., Walsh, M. J., Skelton, R. R., Mattes, M. (2016) Mosaic evolution in cultural frameworks: Skateboard decks and projectile points. In Medoza Straffon, L. (Ed.), *Cultural phylogenetics* (pp. 113-130). Dordrecht, Springer Nature.
- Rali, F., & Griffiths, T. (2010). Words as alleles: Connecting language evolution with Bayesian learners to models of genetic drift. *Proceedings of the Royal Society of London B*, 277, 429-436.
- Salthe, S. (1985). *Evolving hierarchical systems: Their structure and representation*. New York: Columbia University Press.
- Shijulal, N. S., et al. (2010). Networks uncover hidden lexical borrowing in Indo-European language evolution. *Proceedings of the Royal Society B*, doi:10.1098/rspb.2010.1917.
- Simpson, G. G. (1944). *Tempo and mode in evolution*. New York, NY, Columbia University Press.
- Sole, R. V., et al. (2010). Language networks: Their structure, function, and evolution. *Complexity*, 15, 20-26.
- Steels, L. (2015). *The talking heads experiment: Origins of words and meanings*. Berlin, Language Science Press.
- Tamariz, M. & Kirby, S. (2016). The cultural evolution of language. *Current Opinion in Psychology*, 8, 37-43.
- Van Gelderen, E. (2013). The linguistic cycle and the language faculty. *Language and Linguistics Compass*, 7, 233–250.
- Wittmann, H. (1969). The Indo-European drift and the position of Hittite. *International Journal of American Linguistics*, 35, 266-68.
- Zhaxybayeva, O., & Doolittle, W. F. (2011). Lateral gene transfer. *Current Biology*, 21(7), R242–R246.

HUMAN AND PORCINE VOCALIZATIONS AFTER CREATION OF A HUMAN LARYNX

KARI GREEN^{*1}, ROBERT MORRISON², MATTHEW WHEELER³, SCOTT HOLLISTER⁴, and GLENN GREEN^{*1,2}

*Corresponding Author: gegreen@med.umich.edu
 ¹Kresge Hearing Research Institute, University of Michigan, Ann Arbor, MI, U.S.A.
 ²Department of Otolaryngology – Head and Neck Surgery, University of Michigan, Ann Arbor, MI, U.S.A.
 ³Department of Animal Sciences, University of Illinois at Urbana-Champaign, Champaign, IL, U.S.A.
 ⁴Department of Biomedical Engineering, Georgia Institute of Technology, Atlanta, GA, U.S.A

Rare medical conditions such as laryngeal agenesis, tracheal agenesis and complete subglottic stenosis physically prevent speech development in cognitively unimpaired children. We identified twelve such children, ranging from 13 months to 7 years of age. Prior to surgical intervention, all of the children had age-appropriate language comprehension despite physical preclusion of laryngeal sound production. To allow these children to develop speech, we have developed novel surgical techniques to construct or reconstruct a larynx and remove the physical obstacles to speech production. These techniques include laryngeal atresia repair (creation of a larynx from surrounding developmental tissues), laryngotracheal reconstruction (using rib cartilage to create a portion of the airway), slide tracheoplasty (excision and reanastomosis to create a patent airway) and laryngeal reinnervation (surgically bringing in another nerve to function for the recurrent laryngeal nerve). For the most complex anatomic defects, we have developed a bioengineered larynx based on typical human anatomy and suitable for surgical implantation – this bioengineered larynx, to date, has only been used in the piglet.

Aphonic children prior to reconstruction had no laryngeal or oral airflow. Although these children had normal hearing and comprehension of auditory input, they were unable to speak. All of their communication to others was via hand gestures, lip and tongue smacking (percussives using articulators without air stream), and manual signs. The communicative abilities of the children were variable; the most proficient child used upwards of 250 ASL signs (Range of all children studied: 0 - ~250 signs; Median: 25 signs). After reconstruction, these children had self-controlled laryngeal airflow. Despite already having normal

language comprehension, these children developed the ability to vocalize via a progression through the same stepwise development of speech that all typical infants experience in acquiring spoken language (Vihman, 1996). Notably, the presence of lip and tongue smacking in these previously aphonic children was not readily coordinated with the now-normal air stream. However, these post-aphonic children progressed to spoken language much more quickly than typical infants. Most surprisingly, the number of signs and modalities used to communicate prior to the surgery did not correlate with speed of speech acquisition after surgery.

As an additional method to allow aphonic children to access speech, we developed a bioengineered human larynx. Fine resolution computed tomography of an adult male human larynx was segmented to create a stereolithographic file of the human larynx with submillimeter accuracy. The file was modified to allow creation of a soft-tissue matrix and the addition of stem cells. Further modification allowed the insertion of a completely-implantable, externally-controlled sound production source. Larynges were then manufactured in our bioengineering laboratory under well-established conditions for implantable medical devices. Initial testing was in the piglet animal model.

After approval was obtained from the animal use committee, a Yorkshire piglet underwent laryngectomy (excision of its voice box), palatopharyngoplasty (palatal reconstruction into a human configuration) and inferior transposition of the airway. The human bioengineered larynx was then implanted into the piglet in an inferior anatomic position corresponding to adult human laryngeal positioning. Following recovery from anesthesia, the piglet had a human laryngeal configuration with the capacity to produce sounds from an externally controlled Bluetooth device. An iPod was used to control, via Bluetooth, an implanted modified speaker which was a part of the 3D-printed larynx, i.e. all vocalizations were externally controlled by the research team (but not the modification of those vocalizations, equivalent to articulations, which were controlled by the piglet). The distal airway was cannulated for breathing and subglottal pressure was therefore not elevated. Similar to the previously aphonic children after reconstruction, the piglet was able to consistently and repeatedly modify the sounds produced by the human bioengineered larynx with tongue, lip and mouth motion after the reconstruction. The patterns were repetitive and cycling. The repetitive nature was consistent with nonaccidental phonation, similar to the babbling of children, whether developmentally or with reconstructed larynges.

The sine qua non of verbal language is the conscious control of vocal communication. Evolution of spoken language requires transition from involuntary (or vegetative) emotional sound production to conscious sound production; and then from conscious sound production to conscious complex sound production. Bohm et al. (2010) show that after reconstruction, normal human verbal speech development transitions through a series of similar stages: (1) Exclusively vegetative function of the larynx; (2) Voluntary control over vegetative laryngeal functions; (3) Babbling; (4) Simple words; (5) Complex speech. Evolutionary theories of language propose different underlying pathways for the development of complex speech.

152

We provide evidence that language comprehension and speech production can evolve separately. During development of speech production in all previously-aphonic reconstructed children, a distinct and prolonged babbling stage was noted. This implies that infants babble as part of a stepwise, normal process to acquire increased laryngeal control (and not due to a lack of language). Moreover, the experience in the piglet demonstrates that nonhuman mammals with human laryngeal positioning and externally-controlled sound production may be able to modify the sounds in nonaccidental patterns. The temporal separation of the acquisition of the children's ability to produce and comprehend language strongly implies that language skills could develop phylogenetically separately from laryngeal control.

Similar to MacNeilage and Davis (2000), we propose that the stepwise acquisition of spoken language, as exemplified in these previously aphonic children, recapitulates the evolution of language.

- Bohm, L. A., Nelson, M. E., Driver, L. E., & Green, G. E. (2010). Babbling, vegetative function, and language development after cricotracheal resection in aphonic children. *The Laryngoscope*, 120(12), 2494-2497.
- MacNeilage P. F. & Davis, B. L. (2000). Evolution of Speech: The Relation Between Ontogeny and Phylogeny. In Knight, C., Studdert-Kennedy, M., & Hurford, J. (Eds.). *The evolutionary emergence of language: social function and the origins of linguistic form*. (pp. 146 – 160) Cambridge, UK: Cambridge University Press.
- Vihman M. M. (1996). *Phonologic Development: The Origins of Language in the Child*. Cambridge, MA: Blackwell Publishers.

NATURAL SELECTION IN THE MODERN ENGLISH LEXICON

JACK GRIEVE^{*1}

*j.grieve@bham.ac.uk

¹English Language and Linguistics, University of Birmingham, Birmingham, UK

This study tests the degree to which the form and function of 54 newly emerging words predicts their success over time in a multi-billion word corpus of American Twitter collected between 2013 and 2016. A linear model of the change in the relative frequency of each word is computed as a function of word length, part-of-speech, word formation process, and meaning. The analysis finds that the most important predictor of the success of these words is marking a new meaning. Shorter words and words created through standard word formation processes are also found to be more successful over time. These results are interpreted as supporting the theory that natural selection is the driving force behind lexical evolution.

1. Introduction

The evolution of the lexicon is difficult to study because most words, especially new words, are incredibly rare. Observing general patterns of lexical innovation therefore requires access to extremely large and densely sampled corpora of natural language. It has only recently become possible to compile such corpora with the rise in popularity of social media, which deposits huge amounts of informal written language online.

For example, based on an analysis of a 9-billion-word corpus of American Twitter, we identified 54 newly emerging words, which were very uncommon at the end of 2013 but whose usage increased substantially over the course of 2014 (Grieve et al. 2017; Grieve et al. 2018). By extracting a relatively large sample of emerging words, we have been able to make general claims about the form, function, and origin of new word formations in Modern American English. This research has also opened up other lines of research, including investigating the factors that predict whether or not emerging words survive over time.

The goals of this study are therefore to begin to understand the forces that drive the evolution of the modern English lexicon by (1) measuring the degree to which the frequencies of these 54 emerging words have changed on Twitter

between 2014 and 2016 and (2) testing the degree to which a range of factors predict the success of these words.

2. Analysis

In our previous research on lexical innovation, we identified 54 emerging words in a 9 billion word corpus of geocoded American Twitter collected between October 2013 and November 2014 (Grieve et al. 2017, 2018). These 54 emerging words, which are listed in Table 1, grew steadily in popularity on Twitter in 2014 and represent a wide range of different parts-of-speech, semantic domains, and word formation processes.

amirite	cosplay	gainz	lordt	rekt	thotful
baeless	dwk	gmfu	lw	rq	thottin
baeritto	fallback	goalz	тсе	scute	tookah
balayage	famo	idgt	mmmmmmuah	senpai	traphouse
boolin	faved	lfie	mutuals	shordy	unbae
brazy	fhritp	lifestyleeee	nahfr	slayin	waifu
bruuh	figgity	litt	notifs	sqaud	wce
candids	fleek	litty	pcd	tbfh	xans
celfie	fuckboys	lituation	pullout	tfw	yaas

Table 1. The 54 emerging words used in the analysis

To quantify the degree to which the popularity of these words changed between 2014 and 2016, I measured the relative frequency of each word in the November section of the 2014 Twitter corpus, the last month in that corpus, and in a new 9 billion word corpus of geocoded American Twitter from 2016. I then computed the factor by which the relative frequency of each word changed between 2014 and 2016 (i.e. the 2016 relative frequency divided by the 2014 relative frequency). Finally, I calculated the log of this factor so that rises and falls in frequency are measured on comparable scales.

This analysis found that the use of the 54 emerging words changed considerably over time. For example, *unbae* (i.e. 'to break up with') dropped from 148 occurrences per million words (pmw) in November 2014 to 4 occurrences pmw in 2016 (-2.0 logged factor of change), while *brazy* (i.e. 'crazy') rose from

1,745 occurrences pmw in November 2014 to 10,723 occurrences pmw in 2016 (+0.8). Overall, a majority (30/54) words fell in usage over this period.

I then constructed a linear model to predict change in the frequency of the 54 emerging words between 2014 and 2016 as a function of four independent variables: length (in characters), part-of-speech (nominal, verbal, adjectival, other), word formation process (acronym, creative spelling, standard), and whether or not the word marks a new meaning. These four variables were selected as predictors because they provide distinct and basic information about the form and the function of these words. The last three predictors require some explanation.

The part-of-speech variable includes 'Nominal', 'Verbal', and 'Adjectival' categories (rather than 'Noun', 'Verb', and 'Adjective') to allow for multiword units, including phrases represented by acronyms, to be classified. The 'Other' category is small and consists mainly of inserts.

The word formation process variable includes 'Acronyms' and 'Creative Spellings' because these are very common orthographic word formation processes on Twitter, even though they are uncommon in speech. Alternatively, the 'Standard' word formation process category includes all processes that are common in spoken language (e.g. compounds, blends, truncations, derivations, borrowings) (Bauer 1982).

The meaning variable was the most difficult to code. The basic distinction being drawn is between words that have meanings that are not already listed in a standard dictionary (e.g. *balayage*, which refers to a specific hair style) and words that have existing synonyms in Standard English (e.g. *baeless*, which means "to be single"). Creative spellings, which always represent existing words, were coded as marking new meanings only if they were associated with a specific nonstandard meaning of that word (e.g. *gainz*, which specifically refers achieving 'weight gains' through exercise), as opposed creative spelling used for emphasis or other functions (e.g. *yaas*).

The linear model for frequency change as a function of these 4 independent variables was found to be significant (F(8, 45) = 4.28, p < 0.001) with an adjusted r-squared of .33. Most notably, meaning was found to be a relatively strong predictor of emerging word success, with words that mark new meanings being especially successful. In addition, shorter words and words formed using standard word formation processes were also found to be more likely to succeed. Alternatively, part-of-speech was found to have relatively little effect on change in the usage of these words. The complete analysis, including R code and data, is available online (Grieve 2018).

3. Conclusion

Overall, the analysis identified three factors that predict if emerging words will survive on Twitter. The most important of these predictors is whether or not the word marks a new meaning, with words that express new meanings being substantially more likely to survive. This finding suggests that the communicative utility of an emerging word is a strong predictor of its success, at least in this variety of language. If an emerging word fills a semantic gap in the standard lexicon, rather than simply providing a synonym for an existing word, it is more likely to be retained.

Similarly, the analysis found that shorter words are more likely to be succeed. This may be because Twitter places strict limits on text length, creating a communicative context that favours shorter words. Although this effect may therefore be restricted to this particular variety of language, given Zipf's (1936) observation that shorter words are generally more common than longer words, this result may also due to a more general principle of lexical change.

Finally, words generated through standard word formation processes were found to be more likely to succeed than words generated through specialised processes that are largely restricted to written language and computer-mediated communication. This finding suggests that emerging words that are suitable for use across varieties of language are more likely to succeed even on Twitter.

All three of these results support the claim that natural selection is a driving force behind lexical evolution, as Darwin himself first proposed in the *Descent of Man* (2003, 1871), where he wrote that "the survival or preservation of certain favoured words in the struggle for existence is natural selection." In particular, this analysis has found that words that are more useful for communication are more likely to succeed, including words that express unique meanings, that are more efficient, and that can be used across a wide range of communicative contexts.

Whether or not these results hold for other varieties of language is an open question. There are also numerous limitations with the present study, most notably the relatively small number of emerging words under analysis and the limited time frame. Given these issues, the main contributions of this study are (1) to illustrate how lexical evolution can be explored through the quantitative analysis of very large corpora of modern language, (2) to provide a preliminary exploratory analysis of the effect of a variety of factors on the success of emerging words on American Twitter, and (3) to present initial empirical support for a general theory of lexical evolution based on natural selection.

- Bauer, L. (1983). *English Word-Formation*. Cambridge: Cambridge University Press.
- Darwin, C. (2003). The Descent of Man. London: Gibson Square.
- Grieve, J. (2018). R Analysis for Natural Selection in the Modern English Lexicon. Rpubs. http://rpubs.com/jwgrieve/evolangxii
- Grieve, J., Nini, A. & Guo, D. (2017). Analyzing lexical emergence in American English online. *English Language and Linguistics* 21, pp. 99-127.
- Grieve, J., Nini, A. & Guo, D. (2018). Mapping lexical innovation on American social media. Forthcoming in *Journal of English Linguistics*.
- Zipf, G. (1936). The Psychobiology of Language. London: Routledge.

LEARNING TO COMMUNICATE ABOUT CONCEPTUAL HIERARCHIES

Robert X. D. Hawkins^{1*}, Kenny Smith², Michael Franke³, and Noah D. Goodman^{1,4}

*Corresponding Author: rxdh@stanford.edu ¹Department of Psychology, Stanford University, USA ²Centre for Language Evolution, University of Edinburgh, UK ³Department of Linguistics, University of Tübingen, Germany ⁴Department of Computer Science, Stanford University, USA

Natural languages provide speakers with remarkable flexibility in the labels they may use to refer to things (Brown, 1958). In addition to the combinatorial explosion of modifiers afforded by compositionality (Partee, 1995), we have a number of lexicalized nominal terms at our disposal. *Dalmatian, dog,* and *animal* can all truthfully be used to talk about the same Dalmatian at different levels of specificity, with one level of the conceptual hierarchy – the basic-level – generally privileged over the others (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). How these overlapping meanings are learned, and why speakers choose different levels of specificity in different contexts, is increasingly accounted for by probabilistic models of pragmatic language use (e.g. Xu & Tenenbaum, 2007; Graf, Degen, Hawkins, & Goodman, 2016) but there remains a more fundamental evolutionary question: how do multiple levels of reference come to coexist in the lexicon to begin with?

Our hypothesis, motivated both by classic work on concept representations and contemporary work on the selective pressures induced by communication, is that lexicalization of conceptual hierarchies is a function of (1) the structure and statistics of entities in the environment, and (2) the particular contexts in which communication occurs. In particular, we expect hierarchical lexica to form when features can be encoded as predictable clusters and communicative goals require distinctions to be drawn at multiple levels. To test this hypothesis, we designed a repeated reference game in which pairs of participants interactively created an artificial language to communicate with each other about objects in context (e.g. Winters, Kirby, & Smith, 2014; Galantucci & Garrod, 2011).

In this game, participants were paired over the web and placed in a shared environment containing a grid of four objects (Fig. 1A) and a 'chatbox' to send messages from a pre-specified vocabulary of sixteen words (Fig. 1B). On each of ninety trials, one player — the 'speaker' — was privately shown a highlighted target object and allowed to send a single word to help their partner select this object



Figure 1. (A) Example array of elements the matcher must choose from. The target is highlighted for the speaker with a black square. In this *subordinate-required* trial there is a distractor at the same intermediate level (striped circle) as the target, so using any abstract label would be insufficient. (B) Drag-and-drop chat box interface. (C) Hierarchical organization of stimuli.

from the array of distractors. The set of objects was designed to cluster in a fixed three-level hierarchy (Fig. 1C). Distractors could differ from the target at any of the levels, creating three kinds of contexts defined by the finest distinction that had to be drawn. In addition to behavioral trajectories observed over the course of the game, we conducted a post-test to explicitly probe players' lexica. For each word, we asked players to select all objects to which that word applies, allowing us to distinguish between subordinate terms that apply to only one element and abstract terms that apply to multiple elements at an intermediate level of the hierarchy: striped circles, for example.

Critically, we manipulated the statistics of the context in a between-subjects design to test the contribution of communicative relevance to lexicalization. In one condition, all three kinds of context were equally likely, thus providing high diversity in the relevant distinctions that must be drawn. We also ran three control conditions in which a single kind of context dominated, e.g. in the 'subordinate-required' condition, the majority of trials contained distractors that were close neighbors to the target (e.g. Fig. 1A), theoretically requiring speakers to lexicalize a label for each target.

We counted the relative number of subordinate-level terms and abstract terms in the post-test and found that the likelihood of lexicalizing abstractions differed significantly across conditions; in particular, the uniform condition was more likely to give rise to lexica in which multiple levels of reference coexist. This suggests that pragmatic pressures for informativity in a diversity of communicative contexts is instrumental for the lexicalization of hierarchical reference systems. Our separate minds may organize the world into meaningful conceptual hierarchies but our shared language only evolves to reflect this structure when it is communicatively relevant.
Brown, R. (1958). How shall a thing be called? Psychological review, 65(1), 14.

- Galantucci, B., & Garrod, S. (2011). Experimental semiotics: a review. *Frontiers in human neuroscience*, *5*, 11.
- Graf, C., Degen, J., Hawkins, R. X. D., & Goodman, N. D. (2016). Animal, dog, or dalmatian? level of abstraction in nominal referring expressions. In A. Papafragou, D. Grodner, D. Mirman, & J. Trueswell (Eds.), *Proceedings* of the 38th annual conference of the Cognitive Science Society.
- Partee, B. (1995). Lexical semantics and compositionality. In L. R. Gleitman & M. Liberman (Eds.), An invitation to cognitive science, part i: Language (pp. 311–360). Cambridge, MA: MIT Press.
- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive psychology*, 8(3), 382–439.
- Winters, J., Kirby, S., & Smith, K. (2014). Languages adapt to their contextual niche. *Language and Cognition*, 1–35.
- Xu, F., & Tenenbaum, J. B. (2007). Word learning as bayesian inference. *Psychological review*, *114*(2), 245.

LINGUISTIC LAWS IN CHIMPANZEE GESTURAL COMMUNICATION

Heesen, Raphaela ^{*1,2}, Hobaiter, Catherine ³, Ferrer-i-Cancho, Ramon ⁴, and Semple, Stuart ²

*Corresponding Author: raphaela.heesen @unine.ch
¹Institute of Work and Organisational Psychology, University of Neuchâtel, Neuchâtel, Switzerland
²Department of Life Sciences, University of Roehampton, London, UK
³School of Psychology and Neuroscience, University of St Andrews, St Andrews, UK
⁴Complexity & Quantitative Linguistics Lab, Departament de Ciències de la Computació,

⁴Complexity & Quantitative Linguistics Lab, Departament de Ciències de la Computació, Universitat Politècnica de Catalunya, Barcelona, Catalonia

Information theory predicts that signals, as vehicles for information transfer, should be efficient in design (Shannon & Weaver, 1962); they are required to be explicit enough to be understood through a noisy medium, but should otherwise minimize costs for the sender (Bradbury & Vehrencamp, 1998). One way to achieve efficiency in communication systems is through Zipf's law of abbreviation (ZLA), namely to assign shorter signals to more frequent elements and longer signals to less frequent ones (Zipf, 1949).

ZLA arises from "compression" – a mathematical principle promoting efficiency in natural and artificial communications systems (Cover & Thomas, 2006). Compression has been explored rigorously in human language, where ZLA has been shown to arise from a compromise between accuracy and efficiency (Kanwal, Smith, Culbertson, & Kirby, 2017). Beyond ZLA, compression is also linked to Menzerath's law (ML), stating that, the larger the construct, the smaller the size of its constituents (Menzerath, 1954). Both laws have been detected across a wide range of languages (Altmann, 1980; Bentz & Ferrer-i-Cancho, 2016), in non-human communication systems (Ferrer-i-Cancho, Hernández-Fernández, Lusseau, Agoramoorthy, Hsu, et al., 2013) and at the molecular level – in genes, genomes and proteins (Ferrer-i-Cancho, Forns, Hernández-Fernández,

Bel-Enguix, & Baixeries, 2013; Ferrer-i-Cancho, 2013; Li, 2012; Shahzad, Mittenthal, & Caetano-Anollés, 2015).

In animal communication, patterns consistent with ZLA, i.e. a negative relationship between signal magnitude and frequency of use, have been detected in the vocal repertoire of Formosan macaques (Semple, Hsu, & Agoramoorthy, 2010), common marmosets (Ferrer-i-Cancho & Hernández-Fernández, 2013), bats (Luo, Jiang, Liu, Wang, Lin, et al., 2013), and in the non-vocal behavioral repertoire of dolphins (Ferrer-i-Cancho & Lusseau, 2009). Patterns consistent with ML, i.e. an inverse relationship between construct and constituent size, have been reported for vocal sequences of geladas (Gustison, Semple, Ferrer-i-Cancho, & Bergman, 2016) and chimpanzees (Fedurek, Zuberbühler, & Semple, 2017). To assess the true breadth of linguistic laws, however, compression research needs to be expanded to hitherto untested modes of communication. Gestural communication, as a key signaling mode in anthropoid primates (Hobaiter & Byrne, 2011), represents a powerful model to test linguistic laws in non-vocal communication systems beyond humans. Here, we analyzed the duration of gestures and gesture sequences given in the context of social play by wild chimpanzees, living on one community at Budongo Forest Reserve, Uganda. Gestures and sequences were recorded from animals of all age classes.

Our analyses provide the first evidence for compression in the form of linguistic laws in animal gestural communication. Although we did not initially find evidence for ZLA - i.e. an inverse relationship between gesture duration and frequency - in the overall play gesture repertoire, we found agreement with ZLA when analyzing specific gesture subsets listed according to their total duration, D (D is the product of frequency of use of a gesture type and its mean duration). In particular, we found ZLA in subsets of gestures of low D, suggesting that compression has acted by reducing both mean duration and frequency of use. Moreover, we found a negative relationship between number of gestures in a sequence and mean duration of the constituent gestures – consistent with ML. We conclude that coding efficiency is a property shared not only between chimpanzee and human communication systems – but also across diverse taxa and modalities; it is possible that these systems may have converged to similar patterns due to similar evolutionary pressures, namely selection for least effort for sender and receiver.

Acknowledgements

RFC was supported by the grant TIN2017-89244-R from MINECO.

- Altmann, G. (1980). Prolegomena to Menzerath's law. Glottometrika, 2, 1-10.
- Bentz, C., & Ferrer-i-Cancho, R. (2016). Zipf's law of abbreviation as a language universal. Capturing Phylogenetic Algorithms for Linguistics. Presented at the Lorentz Center Workshop, Leiden, October 2015.
- Bradbury, J. W., & Vehrencamp, S. L. (1998). Animal communication. Massachusetts: Sinauer.
- Cover, T., & Thomas, J. (2006). *Elements of information theory* (Vol. 2). New York: Wiley.
- Fedurek, P., Zuberbühler, K., & Semple, S. (2017). Trade-offs in the production of animal vocal sequences: insights from the structure of wild chimpanzee pant hoots. *Frontiers in Zoology*, 14(1), 50. https://doi.org/10.1186/s12983-017-0235-8
- Ferrer-i-Cancho, R., Forns, N., Hernández-Fernández, A., Bel-Enguix, G., & Baixeries, J. (2013). The challenges of statistical patterns of language: The case of Menzerath's law in genomes. *Complexity*, 18. https://doi.org/10.1002/cplx.21429
- Ferrer-i-Cancho, R., & Hernández-Fernández, A. (2013). The failure of the law of brevity in two New World primates. Statistical caveats. *Glottotheory*, 4, 44–55. https://doi.org/10.1524/glot.2013.0004
- Ferrer-i-Cancho, R., Hernández-Fernández, A., Lusseau, D., Agoramoorthy, G., Hsu, M. J., & Semple, S. (2013). Compression as a universal principle of animal behavior. *Cognitive Science*, 37, 1565–1578. https://doi.org/10.1111/cogs.12061
- Ferrer-i-Cancho, R., & Lusseau, D. (2009). Efficient coding in dolphin surface behavioral patterns. *Complexity*, 14, 23–25. https://doi.org/10.1002/cplx.20266
- Gustison, M. L., Semple, S., Ferrer-i-Cancho, R., & Bergman, T. J. (2016). Gelada vocal sequences follow Menzerath's linguistic law. *Proceedings of the National Academy of Sciences*, 113, E2750–E2758. https://doi.org/10.1073/pnas.1522072113
- Hobaiter, C., & Byrne, R. W. (2011). The gestural repertoire of the wild chimpanzee. *Animal Cognition*, *14*(5), 745–767. https://doi.org/10.1007/s10071-011-0409-2
- Kanwal, J., Smith, K., Culbertson, J., & Kirby, S. (2017). Zipf's law of abbreviation and the principle of least effort: Language users optimise a miniature lexicon for efficient communication, 165, 45–52. http://dx.doi.org/10.1016/j.cognition.2017.05.001
- Li, W. (2012). Menzerath's law at the gene-exon level in the human genome. *Complexity*, 17. https://doi.org/10.1002/cplx.20398
- Luo, B., Jiang, T., Liu, Y., Wang, J., Lin, A., Wei, X., & Feng, J. (2013). Brevity is prevalent in bat short-range communication. *Journal of*

Comparative Physiology A, *199*, 325–333. https://doi.org/10.1007/s00359-013-0793-y

- Menzerath, P. (1954). *Die Architektonik des Deutschen Wortschatzes*. Bonn: Dümmler.
- Semple, S., Hsu, M. J., & Agoramoorthy, G. (2010). Efficiency of coding in macaque vocal communication. *Biology Letters*, 6, 469–471. https://doi.org/10.1098/rsbl.2009.1062
- Shahzad, K., Mittenthal, J. E., & Caetano-Anollés, G. (2015). The organization of domains in proteins obeys Menzerath-Altmann's law of language. *BMC Systems Biology*, 9, 44. https://doi.org/10.1186/s12918-015-0192-9
- Shannon, C., & Weaver, W. (1962). *The mathematical theory of communication*. Urbana: The University of Illinois Press.
- Zipf, G. (1949). *Human behaviour and the principle of least effort*. Cambridge: MA: Addison Wesley.

COMMUNICATION OR COGNITION?: A MINIMAL COMPUTATIONAL MODEL OF EVOLUTIONARY INTERACTIONS BETWEEN INDIVIDUAL AND SOCIAL LEARNING

Masahiko Higashi*1, Reiji Suzuki1, and Takaya Arita1

*Corresponding Author: higashi@alife.cs.is.nagoya-u.ac.jp ¹Graduate School of Informatics, Nagoya University, Japan

1. Introduction

Animals including humans adapt to the environment by 2 different mechanisms working on 2 levels, evolution and learning, a population level mechanism and an individual level mechanism, respectively. Learning can further be classified into individual learning and social learning (*IL* and *SL*, hereafter). Our purpose is to understand how complex systems, specifically language emerged through the evolutionary interaction between IL and SL by using an agent-based model. Language is a communication tool but also a cognitive tool while most evolutionary scenarios see it as the former (Reboul, 2015). Indeed, in the brain, utilizing language-related circuits, some form of linguistic knowledge is linked to the external world by producing/perceiving sounds and gestures, and at the same time, is connected to the inner mental world composed of concepts, intentions and reasoning (Berwick et. al., 2013), regardless whether there is a shared computational core (e.g. Strong Minimalist Thesis (Chomsky, 2000)).

We assume the fitness function represents the 2 components of selection. One is the directional component (DC, hereafter) that drives the evolution of fundamental traits underlying the whole linguistic activities. For DC, we do not use a simple unimodal function but a minimal fitness function representing a multimodal fitness landscape with linguistic levels. A linguistic level represents a fitness value expected when communication succeeds, corresponding to the number of traits involved in the trait interaction. The basic idea is as follows. In general, language ability regardless of cognitive or communicative is based on multiple subordinate traits, and its mutual interactions are nonlinear. We can further assume that the more adaptive the trait set is, the more inter-trait

interactions are required. In other words, there is a trade-off between adaptivity of individuals and the strength of the epistatic/nonlinear interactions among phenotypes. We represent this idea by using a simple rugged fitness function (Suzuki & Arita, 2007). The other is the positive frequency-dependent component (PFC, hereafter). It has been pointed out that mutations in grammar cannot be beneficial because a mutant's peers might not understand her (Pinker & Bloom, 1990; Glackin, 2010). To represent the collective adaptivity of language, that is related with not only communicative but also cognitive aspects, we simply define PFC of an individual as the ratio of the others with the same linguistic level. When considering the adaptive evolution, it would be plausible to consider either aspect (cognition or communication) of language evolved under the selection with both components. We hypothesize that the evolution of language as a communication tool depended more on PFC than DC, comparing with that of language as a cognitive tool, while we carefully distinguish between proposing language evolved for A, and proposing language evolved as a system of A (or just as being used in A). We thus can change indirectly the proportion of the both aspects by changing the proportion of both components.

We assume an intergenerationally overlapped population in which each performs IL based on trial-and-error of phenotypic changes and SL based on imitation of phenotypes from the most adaptive individual. We evolved the initial phenotypic values, their plasticity and a SL rate via fitness proportional selection. First, we investigated the behavior of the model without PFC. We found that both IL and SL worked cooperatively, and facilitated the evolution of cognitive aspect, enabling the population to cross a fitness valley repeatedly. In this process, IL enabled an individual to find new adaptive phenotypes through its trial-and-error process. SL enabled the whole population to share such adaptive phenotypes, which brought about the genetic assimilation of acquired phenotypes. SL further facilitated an acquisition of more adaptive phenotypes through IL by increasing the genetic diversity of the population. As the proportion of PFC increased, such a cooperative evolutionary process was more significantly retarded, and the population tended to converge to a lower peak. We can discuss possible scenarios of language evolution based on the results. A basic scenario might be the following. As the population size increased, the chance of interaction among members tended to increase. Therefore, we can assume that it led to an increase in the selection pressure caused by PFC, which tended to decelerate the evolution.

Acknowledgements

This work was supported by MEXT/JSPS KAKENHI Grant Number #4903, JP17H06383.

- Berwick, R. C., Friederici, A. D., Chomsky, N., & Bolhuis, J. J. (2013). Evolution, brain, and the nature of language. *Trends in cognitive sciences*, 17(2), 89-98.
- Chomsky, N. (2000). Minimalist inquiries: The framework. In: Martin, R., Michaels, D., Uriager - eka, J., editors. Step by step: Essays on minimalist syntax in honor of Howard Lasnik (pp. 89–155). Cambridge (Massachusetts): MIT Press.
- Glackin, S. N., (2010). Universal grammar and the Baldwin effect: a hypothesis and some philosophical consequences. *Biology and Philosophy*, 26(2), 201-222.
- Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences*, 13, 707-784.
- Reboul, A. C. (2015). Why language really is not a communication system: a cognitive view of language evolution. *Frontiers in Psychology*, 6, Article ID: 1434.
- Suzuki, R., & Arita, T. (2007). Repeated occurrences of the Baldwin effect can guide evolution on rugged fitness landscapes. In Proceedings of the First IEEE Symposium on Artificial Life (IEEE-Alife '07), 8-14.

WHEN IS A CODE NOT A CODE?

ELIZABETH IRVINE

irvinee@cardiff.ac.uk School of English, Communication and Philosophy, Cardiff University, UK

It is often stated that mind-reading, or the ability to attribute and reason about other's mental states, is a cognitive prerequisite for communication (Clark 1996; Levinson 2006; Wilson & Sperber 2012; Scott-Phillips 2014; Tomasello 2008). One foundational argument for this is that the code model of communication fails to account for how hearers access speaker meaning. Alternatives to the code model focus on identifying speakers' intentions, and so require mind-reading. This argument is very rarely questioned, but it is deeply problematic.

First, the 'classic' code model, where signals are associated with meanings in a simple dictionary-like way, does not even account for all cases of animal communication. If this is true, then it is hardly surprising that it also fails to account for some human communication, (and presumably we would not want to conclude that animal communication also requires mind-reading). A more plausible code model of animal communication is required if it is to be a genuine target.

At least some animal communication uses intentionally produced signals that rely on complex and probabilistic codes, where signal interpretation is sometimes context-sensitive (e.g. Roberts et al. 2012). Importing contextsensitive coding into the code model is crucial, as makes it possible for interpretation of signals to depend on a range of socially and ecologically important inputs. These inputs can include prosody, speaker's line of sight and current activities, non-mentalistic goal recognition, and emotion recognition. Complex decoding processes can also rely on pragmatic presumptions, such that it is worthwhile to pay attention to utterances directed towards oneself (e.g. marked by sustained eye contact), and that utterances are easy to decode. In this 'broad' code model, communication is tied to existing ways of successfully interacting with others, and can be done in a way that does not rely on mindreading.

However, if the claim is that animal communication (code-like) is qualitatively different to human communication (not code-like), given the broad code model, it is not entirely clear what the qualitative difference is. Clearly, there is a lot that the broad code model cannot do. But the broad code model makes it possible to identify speaker meaning across a range of contexts that are likely to be ontogenetically and phylogenetically important, including deictic signals. If this is the case, then it looks like mind-reading is not always necessary for engaging in context-sensitive and pragmatically informed communication.

There is a further problem though. If one views all cognition as essentially coding/decoding information, then *any* model of human communication will be a code model at some level of description. Even if mind-reading turns out to be essential for human communication, an information processing model of mind-reading would be part of a larger code model of communication. Yet mind-reading remains a black box, making it difficult to evaluate exactly when and how it contributes to communication. It is not clear what online and stored inputs are selected and used, and how and which inferences are generated. Further, related to similar debates about animal mind-reading, there are thorny questions about what counts as 'genuine' mind-reading, compared to any other process that uses perceptual cues and stored information to predict and explain behavior. So, even if mind-reading is claimed to be part of a larger code-model of communication, it is currently very unclear what process it refers to.

Following Buckner (2014), it is unlikely that there is a clear-cut answer to this, but he provides an interesting perspective on how to approach mindreading. Buckner follows a Dretskian analysis of representational content and suggests that an agent is more along the continuum of being able to mind-read, or represent mental states, the more able they are to integrate information concerning the candidate mental state in question (and act accordingly), and the better they are able to learn novel cues that mark the presence of that candidate mental state. That is, mind-reading is not an all-or-nothing capacity, but a graded capacity to engage in certain kinds of cognitive processing (and ones in which adult humans do not always excel).

Given this gloss on mind-reading however, it is even less clear what necessity claims about mind-reading in the context of communication amount to, because mind-reading is not a single, monolithic capacity. Instead, it seems likely that different types and levels of mind-reading, at different levels of complexity/flexibility, will be necessary for more or less successful engagement with specific types of communicative acts. In this case though, not only is one of the foundational arguments in favour of the necessity of mind-reading for human communication deeply flawed, but claims about the necessity of mindreading for human communication are at best massively underspecified, and at worst cognitively implausible.

- Buckner, C. (2014). The semantic problem(s) with research on animal mindreading. *Mind & Language*, 29, 566–589.
- Clark, H. (1996). Using language. Cambridge, UK: Cambridge University Press.
- Levinson, S. C. (2006). On the human "interaction engine". In N. J. Enfield, & S. C. Levinson (Eds.), *Roots of human sociality: Culture, cognition and interaction* (pp. 39-69). Oxford: Berg.
- Roberts, A. I., Vick, S.-J., and Buchanan-Smith, H. M. (2012). Usage and comprehension of manual gestures in wild chimpanzees. *Animal Behaviour*, 84, 459-470.

Scott-Phillips, T. C. (2014). Speaking our minds. London: Palgrave Macmillan.

Tomasello, M. (2008). The origins of human communication. MIT Press.

Wilson, D., & Sperber, D. (2012). *Meaning and relevance*. Cambridge, UK: Cambridge University Press.

AGENT MODEL REVEALS THE INFLUENCE OF VOCAL TRACT ANATOMY ON SPEECH DURING ONTOGENY AND GLOSSOGENY

Rick Janssen¹, Scott R. Moisik^{2,1}, and Dan Dediu^{*3,1}

*Corresponding Author: Dan.Dediu@mpi.nl

¹Language and Genetics Department, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

²Linguistics and Multilingual Studies, Nanyang Technological University, Singapore ³Collegium de Lyon, Institut d'Études Avancées, Lyon, France

This study introduces a new model for the investigation of the complex manner in which vocal tract anatomy affects human speech production and may influence language change and evolution. The anatomy of the human vocal tract has long been recognized to play a crucial role in speech production and patterning (Fant, 1971; Ohala, 1983). It imposes discrete relations between articulatory parameters and acoustics (Stevens & Keyser, 2010), with highly nonlinear mappings between them (Stevens, 1968, 1989), and it has been recently suggested that inter-individual and patterned inter-population variation in the anatomy of the vocal tract might play a role in explaining patterns of linguistic diversity (Dediu, Janssen, & Moisik, 2017).

We investigate these complex relationships by instructing a computersimulated agent to learn to reproduce, as well as possible, target speech sounds by controlling the articulators of a detailed 3D geometric model of the human vocal tract based on the VocalTractLab 2.1 (Birkholz, Jackèl, & Kroger, 2006), modified to allow changes in larynx height and hard palate shape. More precisely, the agent minimizes the Euclidean distance (in the F1–F5 formant space) between the target and the produced sounds using a genetic algorithm that optimizes the synaptic weights of a neural network that maps formants to articulatory parameter values¹. Here, we apply this model to two case studies, both using the five-vowel system [a], [æ], [i], [u], and [\ominus], but investigating the effects of variation in different components of the vocal tract.

In the first case study, we revisit the debate concerning the role of *larynx height* in human speech, which has important implications for the evolution of speech

¹Unlike other studies (Guenther, 2006; Kröger, Kannampuzha, & Neuschaefer-Rube, 2009), we focus on the anatomy and not on neuro-developmental effects, resulting in a cognitive architecture deliberately designed to be domain-general and based on well-established machine-learning algorithms; however, our architecture is modular and other learning mechanisms can be plugged-in.

and language (Fitch, Boer, Mathur, & Ghazanfar, 2016; Boë et al., 2013; Lieberman, 2012). Our model generally agrees with the conclusions of Boë, Heim, Honda, and Maeda (2002) that a descended larynx is not a necessary prerequisite for modern human speech, but also highlights that there seems to be an optimal larynx height for vowel production approximating that of a modern human female (supporting De Boer, 2010), with a lower or higher larynx seemingly less well suited to produce the optimally expressive vowel inventories of modern human languages. Our model also allows the investigation of the behavior of the other articulators, and we found that certain articulators (such as the tongue and lips) play an important role in the (imperfect) compensation of larynx height, and that they do not seem to act individually but as components of rather complex articulatory subsystems.

The second case study looked into the more subtle effects of the shape of the hard palate. The hard palate affects the articulatory gestures required to produce a large set of speech sounds, including [1] (Tiede, Boyce, Holland, & Choe, 2004; Zhou, Espy-Wilson, Tiede, & Boyce, 2007; Tiede, Boyce, Espy-Wilson, & Gracco, 20010), sibilants (Weirich & Fuchs, 2011), and high vowels (Mooshammer, Perrier, Geng, & Pape, 2004; Brunner, Fuchs, & Perrier, 2005, 2009). As in the previous case, we first investigated the systematic differences between the acquired and target vowel system in individual agents. However, while larynx height has a relatively large effect, hard palate shape shows very weak and subtle effects. This prompted us to use *iterated learning* in chains of agents to investigate the amplification of anatomical biases (Kirby & Hurford, 2002; Kirby, Dowman, & Griffiths, 2007). More precisely, we tested five chains (in a given chain all agents have the same hard palate anatomy; the five hard palate antomies are: an artifically "low" and an artificially "high" configuration, the "standard" configuration of our model, as well as the palate shapes of two human participants acquired with MRI); each chain was run for 50 generations, starting with the same five target vowels, and we ran 20 independent replications. We found that iterated transmission significantly affects the acoustics and articulation across generations, with most vowels and chains reaching a plateau before 40 generations. We found differences between the five palate shapes (including those of real participants), with a very slight tendency for the vowels to become more similar to each other (see Smith, Tamariz, & Kirby, 2013) on a background of vowel- and hard palate shape-specific effects.

In conclusion, using a detailed geometric model of the vocal tract, coupled with modern machine learning algorithms, allows the precise investigation of subtle effects of anatomical variation on speech (anatomical biases). The study showed human larynx height exerts strong biases that are expressed *ontogenetically* (i.e., during an individual's language acquisition). Weaker biases from the hard palate shape are mainly expressed *glossogenetically* (i.e., across cultural transmission) by an amplification effect.

- Birkholz, P., Jackèl, D., & Kroger, B. J. (2006). Construction and control of a three-dimensional vocal tract model. In *Proceedings of the 2006 ieee international conference on acoustics, speech and signal processing* (Vol. 1).
- Boë, L.-J., Badin, P., Ménard, L., Captier, G., Davis, B., MacNeilage, P., Sawallis, T. R., & Schwartz, J.-L. (2013). Anatomy and control of the developing human vocal tract: A response to lieberman. *Journal of Phonetics*, 41(5), 379–392.
- Boë, L.-J., Heim, J.-L., Honda, K., & Maeda, S. (2002). The potential neandertal vowel space was as large as that of modern humans. *Journal of Phonetics*, 30(3), 465–484.
- Brunner, J., Fuchs, S., & Perrier, P. (2005). The influence of the palate shape on articulatory token-to-token variability. ZAS Papers in Linguistics, 4, 43-67.
- Brunner, J., Fuchs, S., & Perrier, P. (2009). On the relationship between palate shape and articulatory behavior. *The Journal of the Acoustical Society of America*, *125*(6), 3936–3949.
- De Boer, B. (2010). Modelling vocal anatomy's significant effect on speech. *Journal of Evolutionary Psychology*, 8(4), 351–366.
- Dediu, D., Janssen, R., & Moisik, S. R. (2017). Language is not isolated from its wider environment: Vocal tract influences on the evolution of speech and language. *Language & Communication*, 54, 9–20.
- Fant, G. (1971). Acoustic theory of speech production: with calculations based on x-ray studies of russian articulations (Vol. 2). Walter de Gruyter.
- Fitch, W. T., Boer, B. de, Mathur, N., & Ghazanfar, A. A. (2016). Monkey vocal tracts are speech-ready. *Science advances*, 2(12), e1600723.
- Guenther, F. H. (2006). Cortical interactions underlying the production of speech sounds. *Journal of communication disorders*, *39*(5), 350–365.
- Kirby, S., Dowman, M., & Griffiths, T. L. (2007). Innateness and culture in the evolution of language. *Proceedings of the National Academy of Sciences*, 104(12), 5241–5245.
- Kirby, S., & Hurford, J. R. (2002). The emergence of linguistic structure: An overview of the iterated learning model. In *Simulating the evolution of language* (pp. 121–147). Springer.
- Kröger, B. J., Kannampuzha, J., & Neuschaefer-Rube, C. (2009). Towards a neurocomputational model of speech production and perception. *Speech Communication*, 51(9), 793–809.
- Lieberman, P. (2012). Vocal tract anatomy and the neural bases of talking. *Journal* of *Phonetics*, 40(4), 608–622.
- Mooshammer, C., Perrier, P., Geng, C., & Pape, D. (2004). An emma and epg study on token-to-token variability. *Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel*, 36, 47–63.

- Ohala, J. J. (1983). The origin of sound patterns in vocal tract constraints. In *The production of speech* (pp. 189–216). Springer.
- Smith, K., Tamariz, M., & Kirby, S. (2013). Linguistic structure is an evolutionary trade-off between simplicity and expressivity. In *Proceedings of the 35th annual meeting of the cognitive science society (cogsci 2013)* (pp. 1348– 1353).
- Stevens, K. N. (1968). The quantal nature of speech: Evidence from articulatoryacoustic data.
- Stevens, K. N. (1989). On the quantal nature of speech. *Journal of phonetics*, *17*(1), 3–45.
- Stevens, K. N., & Keyser, S. J. (2010). Quantal theory, enhancement and overlap. *Journal of Phonetics*, 38(1), 10–19.
- Tiede, M., Boyce, S. E., Espy-Wilson, C. Y., & Gracco, V. L. (20010). Variability of north american english/r/production in response to palatal perturbation.
- Tiede, M. K., Boyce, S. E., Holland, C. K., & Choe, K. A. (2004). A new taxonomy of American English /r/ using MRI and ultrasound. *The Journal of the Acoustical Society of America*, 115(5), 2633–2634.
- Weirich, M., & Fuchs, S. (2011). Vocal tract morphology can influence speaker specific realisations of phonemic contrasts. In *Proceedings of the international seminar on speech production* (pp. 251–259).
- Zhou, X., Espy-Wilson, C. Y., Tiede, M., & Boyce, S. (2007). An articulatory and acoustic study of "retroflex" and "bunched" American English rhotic sound based on MRI. In *Interspeech* (pp. 54–57).

TRIANGULATING SOUND SYMBOLISM: WHERE TO FIND IT AND HOW TO CREATE IT

Niklas Johansson*

*Corresponding Author: niklas.johansson@ling.lu.se Centre for Languages and Literature, Lund University, Lund, Sweden

Sound symbolism is a universal component of language (Samarin 1978; Blasi et al. 2016), but it can also adapt to language-specific constraints such as phoneme inventories, since different languages often use various, albeit phonetically similar, types of sounds for the same sound symbolic association. It is thus natural to investigate the phenomenon from a bottom-up perspective without any initial assumptions other than that it is a universal, non-arbitrary and flexible association between sound and meaning. However, most previous crosslinguistic studies have been small in scope, and larger-scale studies (Wichmann et al. 2010; Blasi et al. 2016) have not captured many phonetic distinctions important for sound symbolism, e.g. voicing (Ohala 1994; Johansson 2017). Furthermore, experiments have usually focused on matching ready-made sound symbolic words to different stimuli (cf. Ramachandran & Hubbard 2001), rather than investigating how sound symbolic associations develop among language users. The present study attempts to amend these issues by focusing on how sound symbolism operates through a more thorough examination of the phonetic and semantic features involved, both cross-linguistically and experimentally.

First, 344 concepts with claimed universal tendencies (e.g. Swadesh 1971; Goddard and Wierzbicka 2002) were investigated in 245 language families and the phonemes of the linguistic forms were systematically grouped according to phonetically salient parameters to pinpoint the features responsible for each sound symbolic association. 178 statistically significant sound-meaning associations were found based on the standard scores calculated for the occurrence of each sound group in each concept, and their occurrence in all of six geographical macro-areas. In addition, these associations could in turn be correlated with at least 45 out of the 100 items of the regular Swadesh-list (Swadesh 1971), raising several questions about the validity of the list as a tool for establishing genetic relationships. Secondly, 42 macro-concepts were

identified based on cooccurring shared semantic and phonetic features between the significant concepts. Most of these had basic descriptive functions (HARD, SMALL, DARK, UNEVEN, etc.), but also included deictic distinctions and kinship attributes. Furthermore, all identified macro-concepts were found to be grounded in one or several of four types of sound symbolism (cf. Dingemanse 2011; Carling & Johansson 2015): (a) in unimodal imitation, or *onomatopoeia*, based on auditory similarity; (b) in a more indirect and cross-modal type of imitation which is grounded in similarities between the referent and *vocal gestures*, in which the accompanying sounds are only secondarily associated with the meaning (e.g. ROUND and labial sounds which have visually round shapes); (c) in the *frequency code* (Ohala 1994), in which resemblance is based on relation with both indexical and iconic grounds; or in (d) an even weaker type of sound symbolism, based on pure circumstantial, indexical associations, e.g. the association between MOTHER, MILK, BREAST etc. and nasals, since those are the only sounds that infants are able to produce whilst breastfeeding.

Thirdly, four of the confirmed sound symbolic concepts were further investigated through iterated learning experiments (Kirby et al. 2015). Naïve participants were divided into five condition groups which contained ten chains of 15 participants each. They either received no information about the meaning of the word they were about to hear, or that it meant BIG, SMALL, ROUND or POINTY, which created a meaning-bias. The first participant in each chain was then audially presented with a phonetically diverse word and asked to repeat it. Thereafter, the recording of the repeated word was played for the next participant in the same chain. Significant increases of high frequency sounds and sounds produced using the hard palate in the SMALL- and POINTY-conditions, and labial and low frequency in the ROUND-condition, were found after 15 generations. These results further revealed that the continuous SIZE-domain was associated with pitch, while the dichotomous SHAPE-domain was associated with the use of separate tactile and visual vocal gestures.

These findings show considerable cross-linguistic sound symbolic effects on basic vocabulary regardless of language family, and that sound symbolism evidently still is an active part of language. They also illustrate how sound symbolism is based in the human perception of the body and its interaction with the surrounding world which is associated through several types of iconicity with different degrees of directness. Thus, it is likely that sound symbolism has originated as a bootstrapping mechanism (Imai & Kita 2014) and could have had an even more crucial communicative role in earlier stages of human language.

- Blasi, D. E., Wichmann, S., Hammarström, H., Stadler, P. F., & Christiansen, M. H. (2016). Sound-meaning association biases evidenced across thousands of languages. *Proceedings of the National Academy of Sciences*.
- Dingemanse, M. (2011). Ezra Pound among the Mawu: Ideophones and Iconicity in Siwu. In P. Michelucci, O. Fischer and C. Ljungberg (Eds), *Iconicity in Language and Literature*, 15, 39-54. Amsterdam: John Benjamins.
- Goddard, C. & Wierzbicka, A. (eds.) (2002). Meaning and Universal Grammar: Theory and Empirical Findings (2 volumes). Amsterdam/Philadelphia: John Benjamins.
- Imai, M., & Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Phil. Trans. R. Soc. B*, 369(1651).
- Johansson, N. (2017). Tracking linguistic primitives: The phonosemantic realization of fundamental oppositional pairs. In: Matthias Bauer, Angelika Zirker, Olga Fischer & Christina Ljungberg (eds.) *Dimensions of Iconicity*. *Iconicity in Language and Literature 15*. Amsterdam: John Benjamins.
- Johansson, N. & Carling, G. (2015). The De-Iconization and Rebuilding of Iconicity in Spatial Deixis: A Indo-European Case Study. Acta Linguistica Hafniensia: International Journal of Linguistics, 47:1, 4-32.
- Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, 141, 87-102.
- Ohala, J. J. (1994). The frequency code underlies the sound-symbolic use of voice pitch. In L. Hinton, J. Nichols and J.J. Ohala (Eds), *Sound symbolism*, 325-347. Cambridge: Cambridge University Press.
- Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia a window into perception, thought and language. *Journal of consciousness studies*, 8(12), 3-34.
- Samarin, W. (1978) Linguistic adaptation to speech function. In W. Mc Cormack and S. A. Wurm (Eds), Approaches to language: *Anthropological issues*. Mouton: The Hague.
- Swadesh, M. (1971). *The origin and diversification of language*. Edited post mortem by Joel Sherzer. London: Transaction Publishers.
- Wichmann, S., Holman, E. W. & Brown, C. H. (2010). Sound Symbolism in Basic Vocabulary. *Entropy* 12(4), 844-858.

CLUES TO LANGUAGE EVOLUTION FROM A MASSIVE DATASET WITH TYPOLOGY, PHONOLOGY, AND VOCABULARY FROM MANY LANGUAGES

SVERKER JOHANSSON^{*1}

*Corresponding Author: sja@du.se ¹Dalarna University, Falun, Sweden

1. Introduction

A major component in the evolution of language is the evolution of the human language capacity, whatever biological endowments humans have that make us language-ready. But the language capacity is not well understood and is difficult to study directly. Clues may come from biases displayed by humans in language acquisition and language change. Even weak underlying biases can lead to strong patterns in the resulting languages (Smith, 2011). Biases can be studied at the individual level in learning experiments (e.g. Culbertson, 2012, Tamariz et al., 2014), but they can also be inferred at the macro level from patterns in the features of natural languages (e.g. Dediu & Ladd, 2007). Biases can be seen either in the synchronic patterns of language features today, or in the diachronic patterns of transition probabilities between features as languages culturally evolve (e.g. Dunn et al, 2011).

Patterns that reveal biases may be found in any aspect of language, e.g. syntax, morphology, phonology, or lexicon, and may be subtle enough to be discernible only in large samples of languages. This work is an exploratory study across the widest possible set of languages, combining typological, phonological, lexical and phylogenetic data on a significant fraction of the languages of the world, with the goal of mapping any biases that may be present. Both synchronic and diachronic patterns are studied, with the emphasis on the latter.

2. Data set

The following data sources are used:

- **Phylogeny and geography**: Ethnologue (Simons & Fennig 2017); ~7,500 languages.
- **Phonological inventories:** PHOIBLE (Moran & McCloy & Wright 2014); ~1,800 languages.
- Typology: WALS (Dryer & Haspelmath 2013); ~2,500 languages.
- Lexicon (Swadesh lists): Rosetta Project Digital Language Archive (2009); ~1,300 languages.

All four types of data are available for ~300 languages. At least three types are available for ~1,600 languages from 132 different stocks. In order to keep the data set as homogeneous as possible, each type of data has been imported from a single source only. Languages are identified between data sources by their ISO codes.

3. Methods

The language phylogeny from Ethnologue is taken as given in the analysis. For the synchronic analysis, the phylogeny is taken into account in the character statistics by down-weighting multiple "hits" in the same family, in order to control for phylogenetic bias and lineage-specific patterns. Geographic data is also available to control for areal effects. Cross-correlations between different types of characters are analysed for possible patterns.

For the diachronic analysis, the phylogeny together with modern-day character data are used to infer both ancestral character states up the language tree for phonological and typological characters, and transitional probabilities between states (including the probability of characters appearing and disappearing), in a bootstrapping process.

4. Some preliminary results

Well-known typological patterns are reproduced. But correlations between features are observed that go beyond those normally discussed in typology, or those observed by Dunn et al (2011). Interestingly, there are also some modest cross-correlations between grammatical features and phonemes. For example, the presence of aspirated consonants and nasal vowels correlates with certain word-order features, even after controlling for phylogeny.

In the diachronic analysis, there are hints of patterns beyond the obvious one that transition probabilities into common features are larger, but much work remains to be done in the interpretation of these patterns.

- Culbertson, J. (2012) Typological Universals as Reflections of Biased Learning: Evidence from Artificial Language Learning. *Language and Linguistics Compass* 6/5, 310–329
- Dediu, D. & Ladd, D.R. (2007) Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, ASPM and Microcephalin. *Proc Nat Acad Sci 104*, 10944-10949.
- Dryer, Matthew S. & Haspelmath, Martin (eds.) 2013. The World Atlas of Language Structures Online. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info, Accessed on 2017-08-30.)
- Dunn, M., Greenhill, S. J., Levinson, S. C. & Gray, R. D. (2011). Evolved structure of language shows lineage-specific trends in word-order universals. *Nature* 473, 79–82
- Moran, Steven & McCloy, Daniel & Wright, Richard (eds.) 2014. PHOIBLE Online. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://phoible.org, Accessed on 2017-08-30.)
- Roberts, S & Dediu, D & Levinson, S (2012) Detecting differences between the languages of Neanderthals and modern humans. Presented at Evolang 10.
- Rosetta Project Digital Language Archive (2009) <u>http://rosettaproject.org/</u>, <u>https://archive.org/details/rosettaproject</u>
- Simons, Gary F. and Charles D. Fennig (eds.). 2017. Ethnologue: Languages of the World, Twentieth edition. Dallas, Texas: SIL International. Online version: <u>http://www.ethnologue.com</u>.
- Smith, K. (2011) Learning Bias, Cultural Evolution of Language, and the Biological Evolution of the Language Faculty. *Human Biology 83*, 261-278.
- Tamariz, M., Ellison, M., Barr, D.J., & Fay, N. (2014) Cultural selection drives the evolution of human communication systems. *Proc. R. Soc. B: Biological Sciences* 281, DOI: 10.1098/rspb.2014.0488.

ENVIRONMENTAL CONSTRAINTS IN THE EVOLUTION OF SCALAR CONCEPTS: ROAD TO *MOST*

DARIUSZ KALOCIŃSKI

Corresponding Author: d.kalocinski@uw.edu.pl Institute of Philosophy, University of Warsaw, Warsaw, Poland

One of the driving forces of language evolution is the selection of variants that suit the communicative needs of its users. Crucially, fitness of linguistic variants may largely depend on the structure of the environment in which language is learned, transmitted, and used. This hypothesis has gained support in various domains. We apply it in the context of scalar terms with a major focus on quantifiers, such as *most*. Based on a model that combines logic and evolutionary game theory, we argue that such signals might have evolved as stable semantic units through adaptation to general communicative principles and distributional properties of the environment such as normality.

1. Introduction

During the development across various timescales, languages tend to adopt variants exhibiting greater communicative fitness (Christiansen & Chater, 2016a). Crucially, the fitness of linguistic variants may largely depend on the structure of the environment in which language is learned, transmitted, and used. This theoretical stance has gained support in several domains, including spatial descriptions (Levinson, 1996, 2003), color categories (Lindsey & Brown, 2002; Plewczyński et al., 2014), kinship terms (Kemp & Regier, 2012), and constituent order (Christensen, Fusaroli, & Tylén, 2016). Quantifiers, despite their ubiquity in natural language, are less explored from this perspective, with notable exceptions focusing on fuzziness and context-dependence (Pauw & Hilferty, 2012, 2016).

The present paper is an attempt to model the influence of environmental constraints on the evolution of scalar terms, and of proportional quantifiers in particular. *Most* is a paradigmatic example of such a quantifier: it stands out in terms of frequency in natural language corpora (Szymanik & Thorne, 2017) and is present across many languages (Katsos et al., 2016). Our goal is to provide some reasons that *most* might have acquired and sustained its present semantic threshold, commonly associated with 1/2 (Peters & Westerståhl, 2006). Interestingly, our argument heavily relies on the analysis of a much broader class of scalar concepts.

We briefly describe the notion of the scalar concept in Section 2. Section 3 describes a communication game based on the discrimination of two contexts. In

Section 4, we analyse those properties of scalar terms that are important when considering their evolution. We present our main argument in Section 5 and the conclusion in Section. 6.

2. Scalar Concepts

A scalar concept is represented as a threshold dividing a given range of values. This is easily modelled, for example, as a real/rational number dividing an (extended) interval of real/rational numbers. We do not consider the direction of monotonicity here: we assume that a scalar concept comprises all values exceeding the corresponding threshold.

Tentative examples of scalar concepts are *tall* and *most*. *Tall* may be referred to as a first-order concept, as possible thresholds are properties of individual objects, namely different values of height. *Most* may be referred to as a second-order concept, as possible thresholds are not properties of individual objects, but rather properties of sets of individual objects (Barwise & Cooper, 1981). In this particular case, the relevant property is identified with a fraction of cardinalities between corresponding sets. Recall that a typical context for the statement *Most As are B* consists of two finite sets *A*, *B*. The standard meaning of *most* is as follows: *Most As are B* is true iff $|A \cap B| > |B - A|$, equivalently $|A \cap B|/|B| > 1/2$. Substituting other proportions for 1/2 yields other proportional quantifier meanings.

3. Discrimination Game

To put scalar terms into a communicative setting, consider agents using one shared signal, optionally marked with negation. At any given time, an agent associates one particular threshold with the signal, but different agents may associate it with different thresholds. Suppose we have an agent using threshold p and his interlocutor using q. They want to discriminate between two (shared) contexts c and c'.¹ Such a game is a success iff both players can tell the difference between the contexts using their current strategies.² Formally, players receive the payoff 1 iff $min(c, c') \leq min(p, q) \leq max(p, q) < max(c, c')$. Otherwise, the payoff is 0.

The variability of contexts is modelled by a random variable X (environment). Each time a game is to be played, two contexts are drawn from X. When two players meet and they happen to use strategies p and q, the likelihood that they are both capable of discriminating between upcoming contexts is

$$D_{X,p,q} = 2P(X \le \min(p,q))P(X > \max(p,q)).$$

$$\tag{1}$$

¹We use the term "context" in a different way then does Steels (1997). It it is more natural to speak about a quantified sentence evaluated against a context (rather than against object or event happening in a shared context).

²If at least one of the players cannot do this using his current strategy, we view this as a failure: the sender has no incentive to use such a (non-discriminative) signal, while the receiver will not understand which of the contexts is the topic. Our game is a simplified version of that considered in Pauw and Hilferty (2012).

which are the chances of obtaining a pair of random deviates (c, c') satisfying $p \in [c, c')$ and $q \in [c, c')$. Observe that, when p = q, we have $D_{X,p,q} = 2P(X \le p)P(X > p)$. We will use $D_{X,r}$ to denote the value of $D_{X,r,r}$.

4. Look at Evolution

We want to compare various strategies with regard to their communicative fitness under plausible environmental constraints. Our chief assumption about the environment is normality, i.e. $X \sim \mathcal{N}(\mu, \sigma^2)$, for some $\mu \in \mathbb{R}$ and $\sigma > 0$. This is a tentative but fairly realistic assumption, as many properties are known to behave in this way. Our analysis is based on standard tools of evolutionary game theory.³

We consider a large population of agents who are paired randomly to play the discrimination game. For simplicity, we confine our attention to two co-existing strategies, p and q. Let $x_p = 1 - x$ and $x_q = x$ be the fractions of the population using strategies p and q, respectively. The communicative fitness of a p-and a q-individual, given in Equations 2a and 2b, is expressed as the expected payoff she receives from a discrimination game while being randomly paired with other agents (the chances of meeting someone endorsing p and q are 1 - x and x, respectively):

$$E_p(x,X) = (1-x)D_{X,p} + xD_{X,p,q}$$
(2a)

$$E_q(x,X) = (1-x)D_{X,p,q} + xD_{X,q}$$
(2b)

Given strategies p, q and a random variable X, p is said to be evolutionarily stable against q in X if there is $y, 0 < y \le 1$, such that for all x < y, $E_p(x, X) > E_q(x, X)$. We say that p is immune to invasion in X if, for all strategies q such that $q \ne p$, p is evolutionarily stable against q in X.

It is not difficult to see that every strategy is immune to invasion in X.⁴ Hence, once a strategy has dominated the entire population, it cannot be replaced by any other strategy through the invasion of an initially very small number of mutants.

Strategies, although immune to invasion, are not equivalent in terms of their expected payoffs. A crucial property of two (different) strategies p, q is the population threshold x_0 at which $E_p(x_0, X) = E_q(x_0, X)$. A slight modification in the proportions of strategies within a population, $x_0 \pm \varepsilon$, provides an advantage to one of the strategies in terms of communicative fitness.

Figure 1 shows how different strategies are related to each other with regard to population thresholds in the standard normal environment. Observe that an invading strategy q can outperform p = 0 only at levels $x_q > 0.5$, in other words, when invaders comprise more than half of the population. However, when $p \neq 0$, strategies q, which are closer to the mean than p, can outperform it at lower levels,

³See, for example, Easley and Kleinberg (2010).

⁴It follows from the linearity of the expected payoff, and from the fact that, for all $p, q \in \mathbb{R}$ such that $p \neq q, E_p(0, X) = D_{X,p} > D_{X,p,q} = E_q(0, X)$.



Figure 1. Population thresholds f(p,q) for pairs of strategies (p,q) in the environment $X \sim \mathcal{N}(0,1)$ where $f(p,q) = y \Leftrightarrow E_p(y,X) = E_q(y,X)$, for all $p,q \in \mathbb{R}$. If $x_q > f(p,q), q$ outperforms p in terms of communicative fitness. If $x_q < f(p,q)$, the opposite is true.

indicating that even if the invaders are in the minority, they can still spread across the novel semantic threshold successfully. It is interesting that, if strategy q is just a little closer to the mean than is p, the population threshold drops sharply and attains very small values for ps lying further from the mean, indicating that overcoming poor conventions should be relatively easy. However, the closer the prevalent strategy is to the mean, the higher the number of invaders required to compromise its communicative fitness. A major lesson is that closeness to the mean is crucial for immunity as understood in terms of population thresholds. Most importantly, the strategy corresponding to the mean is the most advantageous one in this respect.

By employing replicator dynamics (Hofbauer & Sigmund, 1998), we can obtain insight into how proportions of strategies in a population may change over time. A slight modification to the population threshold for which the rate of growth \dot{x}_p of the *p*-subpopulation equals 0 will lead to the total domination of one strategy. These dynamic variants of population thresholds behave roughly the same as do the ordinary population thresholds described above (see supplementary materials).

5. Road to Most

A typical context for a proportionally quantified sentence (*Most boys are tall*) consists of two sets A, B (the set of boys and the set of tall people, respectively). On many occasions, as in our example, B consists of instances of a scalar concept. Based on our previous considerations, assume that first-order scalar concepts have

developed into thresholds equal to the means of the corresponding normally distributed variables. Let μ be a concept used for the classification of objects with regard to the property $X \sim \mathcal{N}(\mu, \sigma^2)$. Trivially, the probability $P(X > \mu)$ of judging a random deviate of X as member of the concept μ is equal to 1/2.

As in the case of first-order scalar concepts, the variability of contexts is modelled probabilistically. Given a set A consisting of n objects, we envisage that $A \cap B$ is obtained from the sequence X_1, X_2, \ldots, X_n of n independent, identically distributed random variables, $X_i \sim \mathcal{N}(\mu, \sigma^2)$, $i = 1, 2, \ldots, n$. Hence, for an *i*th object of the set A, $P(X_i > \mu) = 1/2$. As a result, $|A \cap B|$ is the number of successes of n independent Bernoulli trials with the probability of success 1/2, $Y_n \sim B(n; 1/2)$.⁵ For simplicity, we do not mention A at all, and assume that each context consists of n objects in total with the number of objects belonging to B governed by $Y_n \sim B(n; 1/2)$. However, what interests us most is the random variable Y_n/n . It turns out that the mean of Y_n/n is 1/2. Based on the normal approximation to binomial distribution, $\mathcal{N}(E[Y/n], Var(Y/n))$ may be treated as a good approximation to Y_n/n , provided n is sufficiently large (a liberal rule of thumb is $np \ge 5$ which, for p = 1/2, yields $n \ge 10$).

Given a normal approximation Z to Y_n/n , the analyses of previous sections applies to proportional quantifiers. Strategies are proportions from [0, 1]. Strategy 1/2 corresponds to *most*. In general, the truth conditions corresponding to a strategy $r \in [0, 1]$ are given by $|A \cap B|/|A| > r$. Hence, the probability that a given context A, B is classified as true with regard to the threshold r is $P(Y_n/n > r)$ which, in turn, is approximated by P(Z > r). All conclusions obtained in previous sections thus remain valid for proportional quantifier strategies, and 1/2 is thus the most advantageous strategy, both in terms of communicative fitness and immunity to invasion. It can also easily invade strategies that are far from 1/2.

6. Conclusions and Perspectives

We have argued that the meaning of *most*, the proportional quantifier occurring so often across so many languages, may be viewed as an adaptation of language to general communicative principles and distributional properties of the environment. Our explanation begins with a normality assumption concerning the properties of individual objects. We show that scalar concepts, when used to refer to such properties, are likely to develop, and sustain, thresholds close the corresponding means. Next, we extend our argument to proportional quantifiers. Assuming the most likely thresholds of first-order scalar concepts, we investigate the probabilistic behaviour of higher-order scalar properties, such as $|A \cap B|/|A|$, which are

⁵Obviously, to obtain a full probabilistic model for generating sets A and B, we should also make some assumptions about A. However, as long as the variables used to obtain A and B are independent, our approach remains valid. We can imagine this in the following way: say our context consists of k objects in total. One variable determines which of the k objects belong to A. The other one selects, independently, which of the objects in A have also the property B.

crucial in comparisons with threshold proportions of proportional quantifiers. We show that, when *B* corresponds to a first-order scalar concept, the probabilistic behaviour of $|A \cap B|/|A|$ is approximately normally distributed with a mean of 1/2. This yields precisely 1/2 as the most favourable threshold for a proportional quantifier—a result that aligns well with the usual interpretations of *most*.

The present findings contribute to theories according to which languages, during their development across various timescales, are driven towards variants exhibiting greater communicative fitness (Christiansen & Chater, 2016a). Our results reveal that proportional quantifiers are no different in this regard than are other, less abstract types of constructions investigated so far. This also leads us to hypothesise that the semantics of quantifiers not mentioned here may be an adaptation of language to the invariant features of the world.

It is interesting that, as exemplified by our analysis, investigating the influence of environmental constraints on quantifier meaning requires a parallel, ecologically valid analysis of such influences on lower-order concepts. This methodological caveat follows directly from the fact that the truth value of quantified sentences depends not only on the meaning of quantifier determiners (denoting higher-order concepts), but also on the meaning of nouns and verb phrases (denoting lower-order concepts) (Barwise & Cooper, 1981).

Communicative pressures are certainly not the only selectional forces that are important for explaining the evolution of meaning. Another important universal pressure stems from our cognitive limitations (Christiansen & Chater, 2008, 2016b). Recent insights at the intersection of formal semantics and cognitive science show that cognitive constraints influence the verification of quantified sentences (Szymanik, 2016). There are also attempts to explain quantifier semantic universals in terms of learnability pressures (Steinert-Threlkeld & Szymanik, 2017). It seems that enriching evolutionary explanations of quantifier meaning with plausible cognitive constraints should make the entire picture more accurate.⁶

Finally, it would be beneficial to validate the formal model proposed in the paper on authentic data obtained from samples collected experimentally or from recorded linguistic usage (corpus studies).

Acknowledgements

The author was supported by the Polish National Science Centre grant number 2015/19/B/HS1/03292.

 $^{^{6}}$ For example, a model by Kalociński, Gierasimczuk, and Mostowski (2015) assumes a tentative simplicity ordering over proportional quantifier meanings based on the size of the denominator, yielding 1/2 as one of the simplest meanings available. See also Kalociński, Mostowski, and Gierasimczuk (2018).

- Barwise, J., & Cooper, R. (1981). Generalized Quantifiers and Natural Language. *Linguistics and Philosophy*, *4*, 159–219.
- Christensen, P., Fusaroli, R., & Tylén, K. (2016). Environmental constraints shaping constituent order in emerging communication systems: Structural iconicity, interactive alignment and conventionalization. *Cognition*, 146, 67–80.
- Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, *31*(05), 489–509.
- Christiansen, M. H., & Chater, N. (2016a). *Creating Language: Integrating Evolution, Acquisition, and Processing.* MIT Press.
- Christiansen, M. H., & Chater, N. (2016b). The Now-or-Never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*, *39*, e62.
- Easley, D., & Kleinberg, J. (2010). *Networks, crowds, and markets: Reasoning about a highly connected world.* New York: Cambridge University Press.
- Hofbauer, J., & Sigmund, K. (1998). *Evolutionary games and population dynamics*. Cambridge University Press.
- Kalociński, D., Gierasimczuk, N., & Mostowski, M. (2015). Quantifier Learning: An Agent-based Coordination Model. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems* (pp. 1853–1854). Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems.
- Kalociński, D., Mostowski, M., & Gierasimczuk, N. (2018). Interactive semantic alignment model: Social influence and local transmission bottleneck. *Journal of Logic, Language and Information*. (to appear)
- Katsos, N., Cummins, C., Ezeizabarrena, M.-J., Gavarró, A., Kuvač Kraljević, J., Hrzica, G., Grohmann, K. K., Skordi, A., López, K. Jensen de, Sundahl, L., Hout, A. van, Hollebrandse, B., Overweg, J., Faber, M., Koert, M. van, Smith, N., Vija, M., Zupping, S., Kunnari, S., Morisseau, T., Rusieshvili, M., Yatsushiro, K., Fengler, A., Varlokosta, S., Konstantzou, K., Farby, S., Guasti, M. T., Vernice, M., Okabe, R., Isobe, M., Crosthwaite, P., Hong, Y., Balčiūnienė, I., Ahmad Nizar, Y. M., Grech, H., Gatt, D., Cheong, W. N., Asbjørnsen, A., Torkildsen, J. v. K., Haman, E., Miękisz, A., Gagarina, N., Puzanova, J., Anđelković, D., Savić, M., Jošić, S., Slančová, D., Kapalková, S., Barberán, T., Özge, D., Hassan, S., Chan, C. Y. H., Okubo, T., Lely, H. van der, Sauerland, U., & Noveck, I. (2016). Cross-linguistic patterns in the acquisition of quantifiers. *Proceedings of the National Academy of Sciences*, *113*(33), 9244–9249.
- Kemp, C., & Regier, T. (2012). Kinship Categories Across Languages Reflect General Communicative Principles. *Science*, 336(6084), 1049–1054.

- Levinson, S. C. (1996). Language and space. *Annual Review of Anthropology*, 25(1), 353–382.
- Levinson, S. C. (2003). Space in language and cognition: Explorations in cognitive diversity (Vol. 5). Cambridge University Press.
- Lindsey, D. T., & Brown, A. M. (2002). Color Naming and the Phototoxic Effects of Sunlight on the Eye. *Psychological Science*, 13(6), 506–512.
- Pauw, S., & Hilferty, J. (2012). The emergence of quantifiers. *Experiments in Cultural Language Evolution*, *3*, 277.
- Pauw, S., & Hilferty, J. (2016). Embodied cognitive semantics for quantification. *Belgian Journal of Linguistics*, 30(1), 251–264.
- Peters, S., & Westerståhl, D. (2006). *Quantifiers in Language and Logic*. Oxford: Oxford University Press.
- Plewczyński, D., Łukasik, M., Kurdej, K., Zubek, J., Rakowski, F., & Rączaszek-Leonardi, J. (2014). Generic Framework for Simulation of Cognitive Systems: A Case Study of Color Category Boundaries. In *Man-Machine Interactions 3* (pp. 385–393). Springer.
- Steels, L. (1997). Constructing and sharing perceptual distinctions. In M. van Someren & G. Widmer (Eds.), Machine Learning: ECML-97: 9th European Conference on Machine Learning Prague, Czech Republic, April 23-25, 1997 Proceedings (pp. 4–13). Berlin, Heidelberg: Springer Berlin Heidelberg. (DOI: 10.1007/3-540-62858-4_68)
- Steinert-Threlkeld, S., & Szymanik, J. (2017). Learnability and Semantic Universals. (unpublished manuscript http://semanticsarchive.net/ Archive/mQ2Y2Y2Z/LearnabilitySemanticUniversals. pdf)
- Szymanik, J. (2016). *Quantifiers and Cognition: Logical and Computational Perspectives* (1 ed.). Springer International Publishing.
- Szymanik, J., & Thorne, C. (2017). Exploring the relation between semantic complexity and quantifier distribution in large corpora. *Language Sciences*, 60, 80–93.

IDENTIFYING LINGUISTIC SELECTION AND INNOVATION WHILE CONTROLLING FOR CULTURAL DRIFT

Andres Karjus^{*1}, Richard A. Blythe^{1,2}, Simon Kirby¹, and Kenny Smith¹

*Corresponding Author: akarjus@exseed.ed.ac.uk ¹Centre for Language Evolution, University of Edinburgh, UK ²School of Physics and Astronomy, University of Edinburgh, UK

1. Introduction

Human languages evolve on a cultural timescale: new linguistic variants are continually being innovated, some variants are selected for, and some are selected against. These processes of selection constitute the engine of language change and, on longer timescales, language evolution. Previous research on lexical competition and selection has generally been based on either limited subsets of natural data (Verkerk, Calude, & Pagel, 2014; Ahern, Newberry, Clark, & Plotkin, 2016; Calude, Miller, & Pagel, 2017) or simulations (Reali & Griffiths, 2010; Blythe & Croft, 2012; Stadler, Blythe, Smith, & Kirby, 2016).

Identifying genuine instances of innovation and selection from natural language corpora on a broad scale is challenging. Language does not exist in a vacuum: discourse topics tend to reflect contemporary social, cultural and political issues. The relative frequency of a word at a given time period is therefore not necessarily an objective measure of its selective fitness at the time, but potentially the result of its related topic(s) being currently more discussed or reported on (cf. also Chelsey & Baayen, 2010; Lijffijt, Sily, & Nevalainen, 2012; Szmrecsanyi, 2016). Here we present the *topical-cultural advection model*, a method which allows us to measure and control for the rise and fall of discourse topics on the frequencies of individual elements (e.g., words).

2. The topical-cultural advection model

The term *advection* is borrowed from physics, denoting the transport of substance (particles) by bulk motion (of fluids), i.e., being carried along by something else (in the linguistic case, words are carried along by topics). We quantify topical fluctuations by measuring change in words associated with the target word by co-occurrence. The advection value of a word in a time period corresponds to the weighted mean of the log frequency changes of its top associated words (its "topic"), with the association scores as weights (building on Santus, Chersoni,

Lenci, Huang, & Blache, 2016; Hamilton, Leskovec, & Jurafsky, 2016). A positive advection value for a given target word therefore indicates its topic is increasing in popularity, and vice versa.

3. Results

We tested the descriptive power of the model in two corpora. In the Corpus of Historical American English (COHA), spanning 20 decades, the (linear) effect of topic changes describes up to 31% of the variance in noun frequency changes.

We also created an artificial corpus in order to test the ability of the topical advection model to identify and correct for false positives when searching for cases of selection. To do this, we synthesized a 26-period corpus, based on two genres, 'spoken' and 'academic', from the Corpus of Contemporary American English: we sampled from the two genres, incrementally increasing one and decreasing the other, simulating a change from academic to spoken style and content. Academic topics are presumably different from spoken topics, but both genres use contemporary English; as such, there should be relatively little evidence of selection in this artificial corpus after controlling for topics. We used a test akin to the Fitness Increment Test (Feder, Kryazhimskiy, & Plotkin, 2014) to identify words undergoing apparent change of frequency, or selection, in this corpus, running the test once without and once with the control for topical advection. In the naive model which did not control for topical fluctuations, 46% of the 5762 persistent nouns in the corpus were selected for (or against), a high rate of false positives. Among the 2624 nouns undergoing putative selection, adjustment for topical fluctuations caused 32% of them to lose the trend, while causing 2% of the nouns to gain a trend (new false positives). In other words, the adjustment using the advection model was capable of eliminating about a third of the false positive cases of selection.

Finally, we tested the model's ability to predict innovation in language. It has been suggested that communicative need (in part) drives vocabulary size (Regier, Carstensen, & Kemp, 2016; Gibson et al., 2017); if so, an increasingly popular topic (i.e. exhibiting positive advection) might attract new words, providing the detailed vocabulary required. We identified 133 successful new nouns entering COHA at the latter half of the 20th century, and found that the advection values of the topics of 55% of the new words were significantly higher at the time they were introduced compared to the previous 10 decades; 38% were around the mean, and 7% below the mean of the advection values of the previous decades. This suggests that the majority of these new words were indeed introduced to talk about increasingly popular topics.

4. Conclusions

We propose the topical-cultural advection model as a reasonable baseline in modelling language change and evolution in general, as a method for removing the topical component in the changing frequencies of linguistic elements in order to better assess their selective fitness, as well as a baseline for considering the fitness of topics in terms of their conductivity to innovations.

- Ahern, C. A., Newberry, M. G., Clark, R., & Plotkin, J. B. (2016). Evolutionary forces in language change. arXiv preprint arXiv:1608.00938.
- Blythe, R. A., & Croft, W. (2012). S-curves and the mechanisms of propagation in language change. *Language*, 88(2), 269–304.
- Calude, A. S., Miller, S., & Pagel, M. (2017). Modelling loanword success a sociolinguistic quantitative study of Mori loanwords in New Zealand English. *Corpus Linguistics and Linguistic Theory*, 0(0).
- Chelsey, P., & Baayen, H. R. (2010). Predicting new words from newer words: Lexical borrowings in French. *Linguistics*, 48(6), 1343–1374.
- Feder, A. F., Kryazhimskiy, S., & Plotkin, J. B. (2014). Identifying signatures of selection in genetic time series. *Genetics*, 196(2), 509–522.
- Gibson, E., Futrell, R., Jara-Ettinger, J., Mahowald, K., Bergen, L., Ratnasingam, S., Gibson, M., Piantadosi, S. T., & Conway, B. R. (2017). Color naming across languages reflects color use. *Proceedings of the National Academy of Sciences*.
- Hamilton, W. L., Leskovec, J., & Jurafsky, D. (2016). Cultural Shift or Linguistic Drift? Comparing Two Computational Measures of Semantic Change. Proceedings of the Conference on Empirical Methods in Natural Language Processing. Conference on Empirical Methods in Natural Language Processing, 2016, 2116–2121.
- Lijffijt, J., Sily, T., & Nevalainen, T. (2012). CEECing the baseline: Lexical stability and significant change in a historical corpus. In *Studies in Variation, Contacts and Change in English* (Vol. 10). Research Unit for Variation, Contacts and Change in English (VARIENG).
- Reali, F., & Griffiths, T. L. (2010). Words as alleles: connecting language evolution with Bayesian learners to models of genetic drift. *Proceedings of the Royal Society B: Biological Sciences*, 277(1680), 429–436.
- Regier, T., Carstensen, A., & Kemp, C. (2016). Languages Support Efficient Communication about the Environment: Words for Snow Revisited. *PLOS ONE*, 11(4), 1–17.
- Santus, E., Chersoni, E., Lenci, A., Huang, C.-R., & Blache, P. (2016). Testing APSyn against Vector Cosine on Similarity Estimation. In Proceedings of the 30th Pacific Asia Conference on Language, Information and Computation, PACLIC 30, Seoul, Korea, October 28 - October 30, 2016.
- Stadler, K., Blythe, R. A., Smith, K., & Kirby, S. (2016). Momentum in Language Change: A Model of Self-Actuating S-shaped Curves. *Language Dynamics* and Change, 6(2), 171–198.

- Szmrecsanyi, B. (2016). About text frequencies in historical linguistics: disentangling environmental and grammatical change. *Corpus Linguistics and Linguistic Theory*, *12*(1), 153–171.
- Verkerk, A., Calude, A. S., & Pagel, M. (2014). Selection in the lexicon. In Evolution of Language: Proceedings of the 10th International Conference (EVOLANG10) (pp. 547–548). World Scientific.

TEMPORAL ADJUSTMENT OF SHORT CALLS DURING VOCAL EXCHANGE IN JAPANESE MACAQUES

NORIKO KATSU 1^{*1,2,3}, KAZUNORI YAMADA 2² KAZUO OKANOYA 3¹, and MASAYUKI NAKAMICHI 4²

*Corresponding Author: katsuno0414@gmail.com
 ¹ Graduate School of Arts and Sciences, University of Tokyo, Tokyo, Japan
 ² Graduate School of Human Sciences, Osaka University, Osaka, Japan
 ³ Japan Society for the Promotion of Science, Tokyo, Japan

1. Introduction

In humans, turn-taking is a common feature during speech. Recent studies have revealed that nonhuman primates (Takahashi et al. 2013) and songbirds (Benichov et al. 2016) avoid overlap in vocal exchange and adjust response latency depending on the call of a partner. These studies also suggest the existence of a dynamic temporal adjustment in vocal exchange among nonhuman animals. Such vocal exchange occurs frequently among individuals with a strong social bond (Snowdon & Cleaveland 1984).

Japanese macaques (*Macaca fuscata*) are known to reply to distance calls (coo calls) of group members after a certain period of latency (Sugiura 2007). Japanese macaques also use short-distance calls, such as grunts, girneys, and short low coos, in affiliative context. These calls are often emitted in a sequence, and are exchanged with a recipient (Katsu et al. 2017). Therefore, investigating vocal exchanges of these short-distance calls used during face-to-face interactions would help to understand social aspects of evolution of human conversation system. In the present study, we examined temporal adjustment in turn-taking of short-distance calls in Japanese macaques.

2. Methods

We conducted behavioral observations using a video camera on 15 adult female Japanese macaques in a free-ranging group in Kyoto, Japan. The group consisted

of 126 members at the time of study. We recorded vocalizations including grunts, girneys, and coo calls of a focal subject facing another individual in the proximity of 5 m (a recipient), using a directional microphone and a digital audio recorder. We also recorded vocalizations from a recipient when the recipient replied within 5 s. We used call bout as a unit for analyses, it is defined as a series of calls emitted in less than 5 s succession. We measured call intervals within a call bout, including inter- and intra-individual intervals.

3. Results and Discussion

We found few overlaps in calls during vocal exchanges in Japanese macaques. Distribution of intra-individual interval of spontaneous calls (i.e., interval between two consecutive calls without a reply from a recipient) showed periodicity of approximately 0.2 s, which is close to the gap between turns in human conversation (0.25 s) (Stivers et al. 2009). There were substantial individual differences in the intervals of spontaneous calls, indicating that the regulation is necessary to avoid overlaps with calls from the recipient.

We then examined if the Japanese macaques adjusted call timing according to the response latency of a partner. We calculated interval between (a) a spontaneous call of a subject and a reply from a recipient, (b) the spontaneous call and a call of the subject after the reply from the recipient, and (c) two spontaneous calls of the subject without replies (Takahashi et al. 2013). The differences between (b) and (c) indicated an amount of temporal adjustment in vocal exchange when compared with that of spontaneous calls, and (a) indicated response latency from a recipient. The analyses revealed that (b)-(c) the amount of temporal adjustment was significantly affected by (a), that is, response latency from a recipient. This finding supports that they adjusted call timing according to the recipient. Japanese macaques are non-vocal learners, however, the present study showed that they have some temporal flexibility in vocalization. A previous study showed that affiliative interactions, such as grooming were more likely to occur when there was vocal exchange (Katsu et al. 2016). Such vocal timing flexibility might play an important role in social coordination in nonhuman primates, and needs further investigation to understand evolution of form and function of turn-taking in human speech.

Acknowledgments

We are grateful for the support provided by Mr. S. Asaba and the staff of the Iwatayama Monkey Park, Japan. Financial support was provided by JSPS KAKENHI Grant #17J10994 to NK and MEXT grant on Innovative Areas #4903, 17H06380 to KO.

- Benichov, J., Benezra, S., Vallentin, D., Globerson, E., Long, M. & Tchernichovski, O. (2016). The forebrain song system mediates predictive call timing in female and male zebra finches. *Current Biology*, 26, 309–318.
- Katsu, N., Yamada, K., & Nakamichi, M. (2016). Function of grunts, girneys and coo calls of Japanese macaques (*Macaca fuscata*) in relation to call usage, age and dominance relationships. *Behaviour*, 153, 125-142.
- Katsu, N., Yamada, K., & Nakamichi, M. (2017). Influence of social interactions with nonmother females on the development of call usage in Japanese macaques. *Animal Behaviour*, 123, 267-276.
- Snowdon, C.T., & Cleveland, J. (1984). "Conversations" among pygmy marmosets. *American Journal of Primatology*, 7, 15-20.
- Stivers, T., Enfield, N.J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G., Rossano, F., de Ruiter, J., Yoon, K-E., & Levinson, S.C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 10587-10592.
- Sugiura, H. (2007). Adjustment of temporal call usage during vocal exchange of coo calls in Japanese macaques. *Ethology*, 113, 528-533.
- Takahashi, D.Y., Narayanan, D.Z., & Ghazanfar, A.A. (2013). Coupled oscillator dynamics of vocal turn-taking in monkeys. *Current Biology*, 23, 2162-2168.
LEARNING IMPLICATIONAL MODELS OF UNIVERSAL GRAMMAR PARAMETERS

Dimitar Kazakov^{*1}, Guido Cordoni², Eyad Algahtani¹, Andrea Ceolin³, Monica A. Irimia⁴, Shin-Sook Kim², Dimitris Michelioudakis², Nina Radkevich², Cristina Guardiano⁴, and Giuseppe Longobardi²

*Corresponding Author: dlk2@york.ac.uk ¹Department of Computer Science, University of York, York, UK ²Department of Language and Linguistic Sciences, University of York, York, UK ³Department of Linguistics, University of Pennsylvania, Philadelphia, US ⁴Dipartimento di Comunicazione ed Economia, UniMORE, Modena, Italy

Abstract

The use of parameters in the description of natural language syntax has to balance between the need to discriminate among (sometimes subtly different) languages, which can be seen as a cross-linguistic version of Chomsky's descriptive adequacy (Chomsky, 1964), and the complexity of the acquisition task that a large number of parameters would imply, which is a problem for explanatory adequacy. Here we first present a novel approach in which machine learning is used to detect hidden dependencies in a table of parameters. The result is a dependency graph in which some of the parameters can be fully predicted from others. These findings can be then subjected to linguistic analysis, which may either refute them by providing typological counterexamples of languages not included in the original dataset, dismiss them on theoretical grounds, or uphold them as tentative empirical laws worth of further study. Machine learning is also used to explore the full sets of parameters that are sufficient to distinguish one historically established language family from others. These results provide a new type of empirical evidence about the historical adequacy of parameter theories.

1. Introduction

In historical linguistics, syntactic parameters can be used as an alternative to phonology and lexicon-based approaches in the attempt to reconstruct phylogenetic trees of languages belonging to one family. Syntactic parameters are also the only type of data that allows to approach the same task for languages belonging to different language families; indeed, by definition, languages from different families do not share lexical features (common etymologies) and the comparison of phonological features has so far been unable to suggest plausible phylogenies

for their apparent lack of sufficient time depth, and for being subject to important secondary contact effects (Longobardi & Guardiano, 2009; Creanza et al., 2015).

Distance and character-based methods (Fitch & Margoliash, 1967; Rannala & Yang, 1996) can be borrowed from population genetics to analyse syntactic parameter data. For each of these approaches, it is important to make explicit any existing dependency between parameters or otherwise the resulting models of the evolution of the languages in question will be biased, since the background assumptions on language typologies will be much looser than the actual conditions constraining possible languages (Bortolussi, Longobardi, Guardiano, & Sgarro, 2011), skewing the probabilistic estimates of historical relatedness. The database developed during the LanGeLin project¹ contains a substantial number of hand-crafted implicational rules of such nature. Here, we add to this body of work by employing Machine Learning techniques to (1) create empirical dependency models between the parameters, and (2) identify the possible groups of parameters whose values are either (2a) shared among all members of a given family, or (2b) are sufficient to separate that one family from all other languages.

The results of (1) allowed us to visualise a very complex network of *possible* dependencies, which have hitherto never been explicitly modelled as a whole. We could then use this empirical data to discuss the previously made choices of parameters and reconsider the existing implicational rules, and make changes where the alternative was deemed more appropriate by the linguistic experts. The results of (2a) have a bearing on hypotheses about the latest common ancestor of all languages in the same family. The results of (2b) can be used as an indicator about possible early evolutionary changes in the history of a given family, which led to its separation as a separate entity (clade). In all cases, the use of machine learning is meant to provide support to historical and evolutionary linguists, rather than replace their expertise and judgement.

2. The Parametric Comparison Method

Parametric theories of generative grammar focus on the problem of a formal and principled theory of grammatical diversity (Chomsky, 1981; Baker, 2001; Roberts, 2012). The basic intuition of parametric approaches is that the majority of observable syntactic differences among languages are derived from a smaller number of more abstract contrasts, drawn from a universal list of discrete, and normally binary, options, called parameters. The relation between observable patterns and the actual syntactic parameters which vary across languages is indirect: syntactic parameters are regarded as abstract differences often responsible for wider typological clusters of surface co-variation, often through an intricate deductive structure. In this sense, the concept of parametric data is not to be simplistically identified

¹LanGeLin ERC Advanced Grant project, 2012–2018.

with that of syntactic pattern: co-varying syntactic properties/patterns are in fact the empirical manifestations of such abstract cognitive structures.

Syntactic parameters are conceived as definable by Universal Grammar (UG), i.e. universally comparable, and set by each learner on the basis of her/his linguistic environment. Open parameters, or any set of more primitive concepts they can derive from (Longobardi, 2005, 2017; Lightfoot, 2017), define a variation space for biologically acquirable grammars, set (a.k.a. *closed*) parameters specify each of these grammars. Thus, the core grammar of every natural language can in principle be represented by a string of binary symbols (Clark & Roberts, 1993), each coding the value of a parameter of UG.

The Parametric Comparison Method (PCM, (Longobardi & Guardiano, 2009)) uses syntactic parameters to study historical relationships among languages. Parameters form a pervasive network of partial implications (Guardiano & Longobardi, 2005; Longobardi & Guardiano, 2009; Longobardi, Guardiano, Silvestri, Boattini, & Ceolin, 2013): one value V_i of parameter A_j , but not the other, may entail the irrelevance of parameter B, whose consequences, i.e. the corresponding surface patterns, become predictable. Under such conditions, B becomes redundant and will not be set by the learner. This rule pattern can be generalised to consider the union of several parameter-value bindings.

An important effect of the pervasiveness of parameter interdependencies is a noticeable downsizing of the space of grammatical variation: according to some preliminary experiments (Bortolussi et al., 2011), the number of possible languages generated from a given set of independent binary parameters is reduced from 10^{18} to 10^{11} when their interdependencies are taken into account. This also crucially implies a substantial reduction of the space of possible languages that a learner has to navigate through when acquiring a language.

3. Learning Dependencies between UG Parameters

Here we adopt an empirical, data-driven approach to the task of identifying parameter dependencies, which has been implemented on our database of 71 languages described through the values of 91 syntactic parameters (see Appendix A) expressing the internal syntax of nominal structures.

We set out to identify parameters whose entire range of values can be fully predicted from the values of other parameters. There is an important difference between previously published work on parameter dependencies and this paper's contribution, which needs to be emphasised: rather than state that, for example, any language in which $P_1 = +$ will have a fully predictable value of P_2 (a fact which we encode as $P_2 = 0$), we seek parameters whose value can be deduced in *all* cases from the values of certain other parameters, e.g. as shown in the hypothetical example in Figure 1. Should such a rule prove to have universal validity, then parameter P_3 would never offer any advantage in distinguishing any two languages, yet it remains a descriptive entity entirely deducible from the other if $P_1 = +$ and $P_2 = -$ then $P_3 = +$ else $P_3 = -$

Figure 1. Parameter dependency model example

parameters.

We process our table of dimensions ($\# lang \times \# param$) with the data mining package WEKA (v.3.6.13) (Hall et al., 2009). More specifically, we take the values of all parameters but one for all languages (i.e. a dataset of size ($\# lang \times$ # param - 1), and learn a decision tree that predicts the value of the remaining parameter from the values of the other parameters. (Typically, only a few are necessary in each case.) This is repeated to produce a decision tree for each of the parameters. The machine learning algorithm used was ID3 (Quinlan, 1986). The algorithm produces a decision tree, in which each leaf corresponds to the value of the modelled parameter for the combination of parameter values listed on the way from the root to that leaf, e.g.: if FGN = - and FGP = + then GCO = + (see Figure 3). Unlike some of the more sophisticated decision tree learning algorithms, such as C4.5 (Quinlan, 1993), no postprocessing of the tree learnt (such as *pruning* (Mitchell, 1997)) takes place, and the tree remains an accurate, exact reflection of the training data. If the combination of parameter values corresponding to one of the leaves of the tree is not observed in the data, the leaf contains the special label 'null' (see the tree predicting GCO in Figure 3). In all other cases, that is, whenever the leaf label is '+', '-' or '0', this is supported by one or more examples (languages) in the data.

```
FGN:
if GCO = 0 then FGN = +
if GCO = + then FGN = -
if GCO = - then FGN = -
GCO:
if FGN = 0 then GCO = null ; never occurs
if FGN = + then GCO = 0
if FGN = - then
    if FGP = 0 then GCO = null; never occurs
    if FGP = + then GCO = +
    if FGP = - then GCO = -
```

Figure 2. Examples of decision trees for parameters FGN and GCO

The decision trees for all parameters were used to produce a dependency graph in which each vertex represents a parameter, and directed edges link the parameters, whose values are needed to predict a given parameter, with the node representing that parameter. For instance, there are edges from both FGN and FGPto GCO, as the decision tree for GCO refers to the values of FGN and FGP. Some of the decision trees are more complex, making use of up to nine separate parameters. The resulting graph is very complex. Therefore, we only present a subset of the graph (see Fig. 3), which only visualises those trees predicting one parameter from the value of one (as in the case of FGN) or two other parameters (e.g. GCO). The fact that some of the rules are missing from this graph is not an issue: for each listed node, all of the incoming edges are present, so that if we know those parameters, we are guaranteed to know the parameter they point to.



Figure 3. Partial dependency graph constructed from implications with up to two antecedents

The interpretation of the graph is straightforward. For instance, looking at its top right corner, one can deduce that for any language in the dataset, it is enough to know the values of parameters EZ3 and PLS in order to know the value of EZ2, and therefore, of EZ1, too. Knowing (the value of) FVP means one also knows DMG and NSD; if one knows both FVP and DNN, the values of DNG, NSD, DSN, DMP and DMG are fully predictable for the given data set. In other words, 7 parameters (FVP, DNN, DNG, NSD, DSN, DMP and DMG are fully predictable for the given data set. In other words, 7 parameters (FVP, DNN, DNG, NSD, DSN, DMP and DMG are fully predictable for the given data set. In other words, 7 parameters (FVP, DNN, DNG, NSD, DSN, DMP and DMG are fully predictable for the given data set. In other words, 7 parameters (FVP, DNN, DNG, NSD, DSN, DMP and DMG) can be reduced to just 2 without any loss of information.

Some of the rules identified by the algorithm are not new, and are already contained in the dataset, as encoded by the implicational system described in Section 1. For instance, the parameter RHM is encoded as 0 when FGP = -, as the value of RHM is fully predictable in those cases. When a decision tree predicting FGP is learned, the result is as follows: if RHM = 0 then FGP = - else FGP = +.

Even the rest of the rules learned are still just empirical findings: they may change with the addition of other examples of languages or their validity may be questioned by linguists on theoretical grounds.

Linguistic analysis of the results is ongoing, and while no part of the results has been accepted as sufficient evidence to dispose of a parameter, implication rules have been revised on the basis of the decision trees learned, as in the case of the parameter PLS. According to our definition, the parameter asks if in a language without grammaticalized Number, a plural marker can also appear outside a nominal phrase, marking a distributive relation between the plural subject and the constituent bearing it. (E.g. PLS = + for Korean, but PLS = - for Japanese.)

Prior to this research, there was an implication rule stating that PLS is neutralised (that is, its value is predictable) for all combinations of CGO and FGN values other than CGO = - and FGN = -. This rule has now been replaced with a new rule stating that PLS is neutralised for all combinations of values of FGM and FGN, except when FGM = + and FGN = -, and the evidence showing that the new rule is consistent with the data came from the tree learned for PLS.

4. Learning Language Family Descriptions

The existing parameters (see Appendix A) have been introduced in order to ensure each language in the database can be uniquely described and separated from the rest on their basis. On a more general level, one could search for the conditions that separate languages from one linguistic family from all others. This is, of course, a classical machine learning task of producing (training) a classifier, which could be used for two purposes, to classify new languages as they are added to the database or to describe the conditions separating one family from the rest. Again, a decision tree can be produced for this purpose. However, it will only contain a very small number of constraints on the parameters that is sufficient for correct classification. Instead, we can adopt another algorithm, namely, Candidate Elimination (CE) (Mitchell, 1997) to learn *all* possible hypotheses (classifiers). This is a classical algorithm for learning in logic, which uses propositional data (i.e. of type $Param_1 = Value_1 \land \cdots \land Param_n = Value_n$) and produces propositional hypotheses, each of which is a conjunction of one or several parameter-value pairs. Each of these hypotheses covers (implies) all positive examples, and does not cover any of the negative (i.e. it is consistent and complete). If no hypotheses of this form and properties can be produced, the result is an empty set of hypotheses. The set of all hypotheses is also known as the *version space* of hypotheses for the given dataset.

While such logic-based approach makes the algorithm rather sensitive to any noise (errors) in the data, here we make the assumption that at this stage of the work, our data is error-free. The output of CE consists of three parts: (1) the set of most specific hypotheses S, i.e. those that cannot be made strictly more specific (by constraining yet another parameter) without becoming incomplete; (2) the set of most general hypotheses G, i.e. those that cannot be made strictly more general without becoming inconsistent, and (3) the rest of the version space, made of hypotheses that are strictly more general than some hypothesis in S, and strictly more specific than some other hypothesis in G.

We applied CE to learning the description of two families of languages, namely, the Romance and the Indo-European (IE), in order to explore the insights it provides. Both families are well established, with the latter subsuming the former. There was a single most specific hypothesis (MSH) for each of the two families (see Table 1). All constraints for the IE family are shared with the Romance family, as expected, while the parameter constraints listed in bold face are specific to the Romance family. This distinction can help guide hypotheses about the last common ancestor of each family, thus providing insight into the evolution of the languages within each family, and the parameters that defined their divergent properties.

Looking at the set \mathcal{G} of most general hypotheses (MGHs) for each family can provide further insight in this direction. While the only MSH in S_{IE} contains 29 parameters (of which 10 zeros, that is, fully predictable), there are numerous MGHs in \mathcal{G}_{IE} that make use of only 2 or 3 parameters, e.g.: (+GSC, -GAL), (-GAL, +PCA), (+FGM, +GSC, -GAL), (+FSN, -XCN, -GUN). A closer look at these parameters reveals that these are particularly useful to delineate boundaries between language families, e.g. -GAL for IE vs. +GAL for Dravidian, Semitic and Uralic languages or +XCN (Dravidian) vs. -XCN (all other families in the database).

5. Discussion

The results reported here show that applying machine learning techniques to the data can reveal previously unknown dependencies between parameters, leading

Indo-European family								
+FGP	+FGM	-FPC	-FGT	+FGN	0GCO	0PLS	+FND	+FSN
-DIN	0FGC	0DBC	-XCN	+GSC	-HMP	+AST	-GCN	0GFN
-GAL	-GUN	-GSI	-ALP	0GST	0GEI	0GNR	0STC	0PMN
+CQU	+PCA							
Romance family								
+FGP	+FGM	-FPC	-FGT	+FGN	0GCO	0PLS	+FND	+FSN
+SGE	+FGG	-CGB	+DGR	0DGP	+CGR	+NSD	-DGD	-DIN
0FGC	0DBC	-XCN	+GSC	+NOE	-HMP	+AST	+FFS	0ADI
-GCN	0GFN	-GAL	-GUN	-EZ1	-EZ2	-EZ3	+GAD	-GFO
0GFS	-GSI	-ALP	0GST	0GEI	0GNR	0STC	-GPC	0PMN
+CQU	+PCA	+PSC	-RHM	+FRC	-NRC	+NOR	0AER	+ARR
-DOR	-NOD	+NOP	+PNP	-NPP	+NOA	+NM2	0FPO	0ACM
-DOA	-NEX	-NCL	0ACL	+TDC	0TNL			

Table 1. Most specific hypotheses for the IE and Romance families

to a potentially significant reduction in the search space of possible languages. Thedata contain more features (i.e. parameters) than data points (i.e. languages), which can make for the generation of spurious rules. The most obvious way to counteract this, adding more languages, comes at a very high cost, as it requires well-trained linguists and an abundance of subtle though typologically wide evidence. One can also use Occam's Razor and limit the search space of possible rules by limiting the number of antecedents in the rule, e.g. to two as we did here. Yet another approach is to collect data selectively for rules of interest, as only a small number of parameters, e.g. 2-3 per language, will be needed to test each rule.

This research could have important implications for the understanding of processes underlying the faculty of language (potentially strengthening the case for UG through strengthening its adequacy as a restrictive typological model and as tool for insightful historical reconstructions), with consequences ranging from models of language acquisition to phylogenetic linguistics, where the syntactic relatedness between two languages may be more adequately measured. However, the approach requires a close collaboration between a machine learning expert, discovering empirical laws in the data, and a linguist who can test their plausibility and theoretical consequences. There is also an open theoretical computational learning challenge here presented by the need to estimate the significance of empirical findings from a given number of examples (languages) with respect to the available range of discriminative features in the dataset.

References

Baker, M. (2001). The atoms of language. New York: Basic Books.

Bortolussi, L., Longobardi, G., Guardiano, C., & Sgarro, A. (2011). How many possible languages are there? In G. Bel-Enguix, V. Dahl, & M. Jiménez-López (Eds.), *Biology, computation and liguistics*. Amsterdam: IOS.

- Chomsky, N. (1964). Current issues in linguistic theory. The Hague: Mouton.
- Chomsky, N. (1981). Lectures on government and binding. Dordrecht: Foris.
- Clark, R., & Roberts, I. (1993). A computational model of language learnability and language change. *Linguistic Inquiry*, 24, 299–345.
- Creanza, N., Ruhlen, M., Pemberton, T. J., Rosenberg, N. A., Feldman, M. W., & Ramachandran, S. (2015). A comparison of worldwide phonemic and genetic variation in human populations. *Proceedings of the National Academy* of Sciences, 112(5), 1265–1272.
- Fitch, W. M., & Margoliash, E. (1967). Construction of phylogenetic trees. Science, 155(3760), 279–284.
- Guardiano, C., & Longobardi, G. (2005). Parametric comparison and language taxonomy. In M. Batllori, M. L. Hernanz, C. Picallo, & F. Roca (Eds.), *Grammaticalization and parametric variation* (pp. 149–174). Oxford: OUP.
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software. *ACM SIGKDD Explor. Newsl.*, *11*, 149–174.
- Lightfoot, D. W. (2017). Discovering new variable properties without parameters. *Ling. Anal.*, *41*. (Spec. ed.: Parameters: What are they? Where are they?)
- Longobardi, G. (2005). A minimalist program for parametric linguistics? In H. Broekhuis, N. Corver, M. Huybregts, U. Kleinhenz, & J. Koster (Eds.), *Organizing grammar: Linguistic studies*. Berlin/NY: Mouton de Gruyter.
- Longobardi, G. (2017). Principles, parameters, and schemata: A radically underspecified UG. *Ling. Anal.*, *41*. (Spec. ed.: Parameters: What are they?) Where are they?)
- Longobardi, G., & Guardiano, C. (2009). Evidence for syntax as a signal of historical relatedness. *Lingua*, 119(11).
- Longobardi, G., Guardiano, C., Silvestri, G., Boattini, A., & Ceolin, A. (2013). Toward a syntactic phylogeny of modern Indo-European languages. *Journal of Historical Linguistics*, *3*(1), 122–152.
- Mitchell, T. (1997). Machine learning. MIT.
- Quinlan, R. (1986). Induction of decision trees. Machine Learning, 1(1), 81-106.
- Quinlan, R. (1993). *C4.5: Programs for machine learning*. San Matteo, CA: Morgan Kaufmann Publ.
- Rannala, B., & Yang, Z. (1996). Probability distribution of molecular evolutionary trees: A new method of phylogenetic inference. *Journal of Molecular Evolution*, 43(3), 304–311.
- Roberts, I. (2012). On the nature of syntactic parameters: a programme for research. In C. Galves, S. Cyrino, R. Lopes, F. Sandalo, & J. Avelar (Eds.), *Parameter theory and language change* (pp. 319–334). Oxford: OUP.

FGP	gramm. person	GSI	grammaticalised inalienability
FGM	gramm. Case	ALP	alienable possession
FPC	gramm. perception	GST	grammaticalised Genitive
FGT	gramm. temporality	GEI	Genitive inversion
FGN	gramm. number	GNR	non-referential head marking
GCO	gramm. collective number	STC	structured cardinals
PLS	plurality spreading	GPC	gender polarity cardinals
FND	number in D	PMN	personal marking on numerals
FSN	feature spread to N	CQU	cardinal quantifiers
FNN	number on N	PCA	number spread through cardinal adjectives
SGE	semantic gender	PSC	number spread from cardinal quantifiers
FGG	gramm. gender	RHM	Head-marking on Rel
CGB	unbounded sg N	FRC	verbal relative clauses
DGR	gramm. amount	NRC	nominalised relative clause
DGP	gramm. text anaphora	NOR	NP over verbal relative clauses/
CGR	strong amount		adpositional genitives
NSD	strong person	AER	relative extrap.
FVP	variable person	ARR	free reduced rel
DGD	gramm. distality	DOR	def on relatives
DPQ	free null partitive Q	NOD	NP over D
DCN	article-checking N	NOP	NP over non-genitive arguments
DNN	null-N-licensing art	PNP	P over complement
DIN	D-controlled infl. on N	NPP	N-raising with obl. pied-piping
FGC	gramm. classifier	NGO	N over GenO
DBC	strong classifier	NOA	N over As
XCN	conjugated nouns	NM2	N over M2 As
GSC	c-selection	NM1	N over M1 As
NOE	N over ext. arg.	EAF	fronted high As
HMP	NP-heading modifier	NON	N over numerals
AST	structured APs	FPO	feature spread to genitive postpositions
FFS	feature spread to struct. APs	ACM	class MOD
ADI	D-controlled infl. on A	DOA	def on all +N
DMP	def matching pron. poss.	NEX	gramm. expletive article
DMG	def matching genitives	NCL	clitic poss.
GCN	Poss°-checking N	PDC	article-checking poss.
GFN	Gen-feature spread to Poss°	ACL	enclitic poss. on As
GAL	Dependent Case in NP	APO	adjectival poss.
GUN	uniform Gen	WAP	wackernagel adjectival poss.
EZ1	generalized linker	AGE	adjectival Gen
EZ2	non-clausal linker	OPK	obligatory possessive with kinship nouns
EZ3	non-genitive linker	TSP	split deictic demonstratives
GAD	adpositional Gen	TSD	split demonstratives
GFO	GenO	TAD	adjectival demonstratives
PGO	partial GenO	TDC	article-checking demonstratives
GES	GenS	TLC	Loc-checking demonstratives

Appendix A: List of Parameters

THE ROLE OF TEACHING IN ITERATED LANGUAGE TRANSMISSION

Vera Kempe^{*1}, Kamil Cichon¹, Monica Tamariz² and Nicolas Gauvrit³

*Corresponding Author: v.kempe@abertay.ac.uk
¹Division of Psychology, Abertay University, Dundee, Scotland, UK
²Department of Psychology, Heriot Watt University, Edinburgh, Scotland, UK
³Université Paris 8 & École Pratique des Hautes Études, Paris, France

Iterated language learning studies have shown that constraints on learning and communication drive emergence of linguistic structure (e.g. Kirby, Tamariz, Cornish & Smith; 2015). However, language is typically passed on by experts who select and modify language input in ways that facilitates learning by novices. Research on first language acquisition has shown that the kind of input enhancement that characterises child-directed speech is beneficial for language learning (e.g. Eaves, Feldman, Griffiths & Shafto, 2016). This suggests that in addition to its other functions such as expressing positive affect, promoting emotional bonding and regulating infant arousal (Singh, Morgan & Best, 2002; Uther, Knoll & Burnham, 2007), this speech register can be considered a form of teaching (Kline, 2015). While the adaptive value of teaching for the transmission of complex cultural traits, and the role of language as a medium of teaching, have recently received increased attention (e.g. Csibra & Gergely, 2009; Laland, 2017), detailed explorations of the effect of teaching on the transmission of language itself are lacking. Here we use iterated teaching of a simple signalling system to explore how language change that emerges during transmission is affected by experts' intention to teach the language to novices.

We compared six transmission chains, of ten generations each, in which a participant's language output after learning was saved and then transmitted to the next participant by the experimenter (*simple transmission condition*) to six transmission chains in which the participants themselves were asked to teach the next participant in the chain (*teaching condition*). Participants learned an "alien" language consisting of binary auditory sequences of high and low tones produced using two programmable buzzers. At the outset, eight randomly

generated 6 or 8-bit binary auditory sequences were associated with meanings representing *kiki*- and *bouba*-type objects differing in size (large vs. small), shape (spiky vs. round) and brightness (dark vs. light) while still retaining individuating features. Signal-meaning pairs were presented twice in incremental fashion. In the simple transmission condition, only six signalmeaning pairs were presented and homophones were removed, to prevent languages from degenerating into ambiguous systems. In the teaching conditions, participants saw all eight pairs and were asked to refrain from speaking while demonstrating the language to the next learner.

The results showed that while in the simple transmission condition individual signals increased in length and languages degenerated slightly despite the homophone filter, in the teaching condition signals retained their length and languages remained expressive by retaining unique signals for the meanings. Similarity of signals within languages increased in both conditions suggesting that learners adopted a set of constraints governing signal production. Algorithmic complexity of individual signals did not decrease and iconicity and compositional structure did not emerge in either condition, presumably due to lack of communicative pressure. Crucially, however, transmission fidelity improved only in the simple transmission condition but not in the teaching condition suggesting that teachers introduced many innovations. We analysed to what extent these teacher innovations were beneficial for learners, and found that while signals did not become easier for learners to reproduce over the course of transmission, they became easier to repeat immediately, both on the first and second attempts. Intriguingly, further analyses revealed that most teacher innovations were triggered by learner repetition errors, and that when learners had erred on their first repetition attempt, they were more likely to correctly repeat the signal on the second attempt the more the teachers' second signals were modified to resemble their first erroneous repetitions. Thus, learner repetition ability was aided by teacher accommodation to learner responses.

These findings suggest that iterated language teaching can affect some functional aspects of language such as expressivity and ease of repetition. We suggest that such teaching effects can arise through two processes: On the one hand, teaching involves spontaneous modifications, i.e. input pre-selection and enhancement, based on teachers' assumptions about learners' knowledge. On the other hand, teaching accommodates direct learner feedback as teachers' productions come to resemble learner responses during episodes of didactic and communicative interaction. Further research should explore the differential effects of these component processes of teaching during language transmission.

References

- Csibra, G. & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Science*, 13, 148-153.
- Eaves Jr, B. S., Feldman, N. H., Griffiths, T. L., & Shafto, P. (2016). Infantdirected speech is consistent with teaching. *Psychological Review*, 123(6), 758.
- Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, 141, 87-102.
- Kline, M. A. (2015). How to learn about teaching: An evolutionary framework for the study of teaching behaviour in humans and other animals. *Behavioural and Brain Sciences*, *38*, E31.
- Laland, K. N. (2017). The origins of language in teaching. *Psychonomic Bulletin* & *Review*, 24(1), 225-231.
- Singh, L., Morgan, J. L., & Best, C. T. (2002). Infants' listening preferences: Baby talk or happy talk? *Infancy*, 3(3), 365-394.
- Uther, M., Knoll, M., & Burnham, D. (2007, January). Do you speak E-NG-L-I-SH? A comparison of foreigner- and infant-directed speech. *Speech Communication*, 49(1), 2–7.

INCREMENTAL WORD PROCESSING HELPS SHAPE THE LEXICON

Adam King*1 and Andrew Wedel1

*Corresponding Author: adamking@email.arizona.edu ¹Department of Linguistics, University of Arizona

One of the most widely attested patterns in human language is that more frequent words tend to be shorter and less frequent words tend to be longer (Zipf, 1935). Further work showed that average probability of words in context (Piantadosi, Tily, & Gibson, 2011) is a better predictor of word length, updating Zipf's original observation. This pattern has been proposed to arise as result of pressures for languages to evolve to become more efficient communication systems: listeners need more information to accurately identify less probable words, and on average, a greater number of sounds in a word provides more information to a listener for word recognition. However, not all sounds in the word are uniformly informative for word recognition.

Listeners process spoken words incrementally, continually updating hypotheses about the identity of the word as they perceive each sound in sequence (Mc-Clelland & Elman, 1986; Norris & McQueen, 2008). As a consequence, earlier sounds in the word contribute more information on average to word recognition than later sounds, because they can, on average, rule out more possible alternatives (see Fig. 1). Here, we show that word that are less contextually probable are more likely to begin with highly informative sequences of sounds. Specifically, less predictable words are more likely to begin with sounds that rapidly distinguish the word from others in the lexicon. This is consistent with previous evidence that the lexicon is under pressure to evolve to serve as an efficient code, and further that as well as affecting the length of words, this pressure can affect the distribution of the sounds that make up the word with respect to the rest of the lexicon.

Methods and Results

We collected word frequencies in 5 languages: English, Dutch and German (Baayen, Piepenbrock, & Gulikers, 1995) and Japanese and Arabic (CallHome Corpus). We restricted our analysis to unaffixed words to avoid any potential confound with the effect of suffixes or prefixes on word processing. Following Son and Pols (2003) and Cohen-Priva (2015), we calculated the information of each phonetic segment of these words, as the -log probability of the segment, given the previous segments in the word, e.g. for the [f] in sphinx, the -log frequency of



Figure 1. The average information provided by each segment position for English and Japanese words. Less frequent words contain more informative segments and this effect is much stronger early.

words with initial [sf] divided by words with the initial segment [s]. We calculated this in two ways: one including a measure of word frequency and one without.

We constructed a mixed-effect linear model to predict segmental information given position in the word and the word's frequency, with a random intercept for each word. Less frequent words contain significantly more informative sounds, independent of length for all languages. There was a significant interaction between position in word and word frequency, indicating that the effect of word frequency on segmental information is primarily found at the beginning of the word. We then fit a linear regression line to segmental information for each word. We found that less frequent words begin with significantly more informative segments. As a further control, we then compared the per-word regressions against those of a novel lexicon in which the order of all segments was reversed (see Fig. 2). We found that there was a much stronger effect of frequency in the unmodified lexicon.

Overall, our findings show that less expected contain more information in their sounds and this extra information is preferentially early. These results suggest that the lexicon has evolved to a state where the lexicon is partially optimized for listeners incremental processing of words.



Figure 2. The initial information for words in English and Japanese, indicating this preference for early information is a product of the linear order of segments.

References

- Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). The celex lexical database (release 2): Linguistic data consortium. *University of Pennsylvania, Philadelphia, PA, USA*.
- Cohen-Priva, U. (2015). Informativity affects consonant duration and deletion rates. *Laboratory Phonology*, 6(2), 243–278.
- Dick Goldhahn, T. E. . U. Q. (2012). Building large monolingual dictionaries at the leipzig corpora collection: From 100 to 200 languages. In *Proceedings* of the 8th international language ressources and evaluation (lrec'12).
- McClelland, J. L., & Elman, J. L. (1986). The trace model of speech perception. *Cognitive psychology*, 18(1), 1–86.
- Norris, D., & McQueen, J. M. (2008). Shortlist b: a bayesian model of continuous speech recognition. *Psychological review*, *115*(2), 357.
- Piantadosi, S. T., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences*, 108(9), 3526–3529.
- Son, R. van, & Pols, L. C. (2003). How efficient is speech. In Proceedings of the institute of phonetic sciences (Vol. 25, pp. 171–184).
- Zipf, G. K. (1935). The psycho-biology of language.

THE ORIGINS OF WORD ORDER UNIVERSALS: EVIDENCE FROM CORPUS STATISTICS AND SILENT GESTURE

SIMON KIRBY, JENNIFER CULBERTSON, AND MARIEKE SCHOUWSTRA

Centre for Language Evolution, University of Edinburgh, UK simon@ling.ed.ac.uk, jennifer.culbertson@ed.ac.uk, marieke.schouwstra@ed.ac.uk

Why are some patterns of noun phrase order (e.g., N-Adj-Num-Dem, 'houses big five these') much more common than others (e.g., N-Dem-Num-Adj, 'houses these five big')? An intriguing possibility is that this distribution emerges in response to a general cognitive bias favouring a transparent relationship between conceptual structure and linear order – an isomorphism bias. In conceptual structure, Adj is closest to N, then Num, then Dem (Fig. 1A). Linear orders that can be read off this nested structure are isomorphic, and these orders are the most common cross-linguistically. Previous experimental work has found an isomorphism in both traditional artificial language learning (Culbertson & Adger, 2014), and silent gesture experiments (Culbertson et al., 2016). For example, Culbertson et al. (2016) asked non-signing English speakers to communicate simple pictures using only gesture. Items were groups of 4 or 5 (Num) objects (N), either spotted or striped (Adj), in a proximal or distal location (Dem). Participants spontaneously improvised isomorphic gestures that did not reflect their native language. These results suggest that an isomorphism bias is present, however they leave open the *origin* of this bias. In particular, they cannot tell us the origins of the conceptual structure that word order is isomorphic to. Here we show that the conceptual structure of the noun phrase is learnable by observing simple statistics about objects in the world.

Intuitively, our proposal is this: properties (~Adj) are more inherent to objects than numerosities (~Num), and location or discourse status (~Dem) is generally not an inherent feature of objects (cf. Rijkhoff 2002). More precisely, we quantify this notion using *point-wise mutual information* (Fig. 1B), a measure of the strength of association between pairs of elements. Using linguistic corpora as a proxy for the world, we can measure average pmi

between nouns and each class of modifier (within a certain window). By averaging over all pairs of, for example, numerals and nouns, we can get an overall measure of how inherent numerosity is to objects. Large corpora of 8 different languages (English, German, French, Spanish, Italian, Portuguese, Chinese, Arabic), plus all English child-directed corpora in CHILDES, confirm our intuition: average pmi of Adj and Num were higher than Dem in all cases, and average pmi of Adj was higher than Num in all but the Portuguese corpus (e.g., Fig. 1C). The conceptual structure in Fig. 1A is thus likely learnable from properties of the world. Interestingly, our measure of inherentness derives a new prediction about sub-classes of adjectives (cf. Martin 1969, Bouchard 2002). Specifically, the corpus results suggest that *size* may not be as inherent to objects as *texture/color* (which pattern together, see Fig. 1C). This predicts that size adjectives may show a weaker isomorphism bias.



Fig 1. A: conceptual structure. B: point-wise mutual information. C: average pmi across modifier types, and adjective subclasses in English (Brown corpus).

To test this, we adapt Culbertson et al.'s (2016) silent gesture experiment, swapping *big* or *small* for the original adjectives. We replicate the original findings for isomorphic order of Num and Dem, but as predicted, the preference for isomorphic order of Adj relative to Num (and to some extent even Dem) is weakened. In other words, isomorphism can be modulated by average pmi, a measure of relative inherentness. This supports the claim that the isomorphism bias taps into a conceptual structure reflecting statistical properties of the world.

In summary, the underlying conceptual structure of the noun phrase—which shapes the typological distribution of orders in this domain—is learnable from observing the statistical properties of the world: it reflects different strengths of associations between objects and their properties, numerosities and locations.

References

- Bouchard, D. (2002). Adjectives, number and interfaces: Why languages vary. Elsevier, Amsterdam.
- Culbertson, J., & Adger, D. (2014). Language learners privilege structured meaning over surface frequency. *Proceedings of the National Academy of Sciences*, 111(16), 5824-5847.
- Culbertson, J., Kirby, S., & Schouwstra, M. (2016). Word order universals reflect cognitive biases: Evidence from silent gesture. In S.G. Roberts, C. Cuskley, L. McCrohon, L. Barcelo-Coblijn, O. Feher & T. Verhoef (eds.) *The Evolution of Language: proceedings of the 11th International Conference*. doi:10.17617/2.2248195
- Martin, J. E. (1969). Semantic determinants of preferred adjective order. Journal of Verbal Learning and Verbal Behavior, 8(6):697–704.
- Rijkhoff, J. (2002). The noun phrase. Oxford University Press, USA.

CONSTITUENT ORDER IN IMPROVISED GESTURE REFLECTS SPEAKER PERSPECTIVE

Fiona Kirton*1, Marieke Schouwstra1, Jenny Culbertson1, Kenny Smith1, and Simon Kirby1

*Corresponding Author: s1354886@sms.ed.ac.uk ¹Centre for Language Evolution, University of Edinburgh, UK

The majority of languages with a dominant word order use either SOV or SVO (Dryer, 2013). The improvised gesture paradigm, in which participants use only gesture to convey information, is increasingly being used to investigate this asymmetry. In one of the earliest studies of this kind, Goldin-Meadow et al. (2008) claimed that Agent-Patient-Action, (here represented as APV but typically equated with SOV), reflects the 'natural' order of elements in improvised gesture. Other authors argue that APV is the natural order only for some types of event and that constituent order in improvised gesture reflects certain properties of an event, such as its temporal structure (Christensen et al., 2016) or the semantic relation between entities and actions (Schouwstra & Swart, 2014). Meir et al. (2017) suggest that gesture order is conditioned on saliency: human entities are more cognitively salient than inanimate entities and are therefore expressed first. Here we investigate the role of saliency in more detail. We present evidence that manipulating the visual saliency of the agent can influence the relative order of other constituents.

Twenty-eight participants were shown pictures of scenes in which a human agent performed an action on an inanimate patient, for example, a man kicking a large potted plant (Fig. 1(a)). They were instructed to describe each scene using only improvised gesture and no speech. Participants were randomly assigned to one of two conditions: the 'generic' condition in which agents represented generic humans such as a man or a woman, or the 'character' condition where more visually salient agents were presented, such as a pirate or a punk. Patients were inanimate objects of a similar size to the agents and were depicted in a state of falling as a result of the action.

We found that in the subset of trials where the agent, patient and action were expressed exactly once, the predominant order in the character condition was AVP; in the generic condition the majority order was APV (Fig. 1(b)). However, looking across all trials, we found that participants were significantly more likely to omit the agent in the generic condition (62% of trials) compared with the character condition (17%) (p<0.001). This suggests that participants in the



generic condition may have attended more to the patient than to the agent.

Figure 1. (a) An example event. (b) Proportions of three-element orders. Includes trials where the agent, patient and action were each expressed once.

Looking at the relative ordering of the patient, which was expressed in a majority of trials in both conditions, and the action, which was expressed on all trials, we found a significant effect of condition such that participants in the generic condition were more likely to express the patient before the action (78%) compared with the character condition (47%) (p<0.05).

Based on these findings, we propose that structural choices in improvised gesture may reflect different ways of mentally construing events. In the generic condition, participants may tend to focus on the role of the patient and therefore construe the action as something that affects, or is experienced by, the patient. Hence, it is more natural to introduce the patient before the action is performed. In the character condition, participants may focus on the more visually salient agent, and so construe the event from the agent perspective: the action is represented as something performed by the agent, and directed towards to the patient. Hence, the agent and action more naturally precede the patient.

Previous studies (see references above) suggest that APV is the natural order only for certain types of event rather than the default structuring principle. Our results add an additional layer to that argument: naturalness as it relates to constituent order is conditioned on the properties of the event, but is also mediated by the way in which speakers construe an event. This, in turn, is influenced by the visual saliency of the participants in the event. The finding that visual attention affects structural choices is not new (e.g., Gleitman et al., 2007), however this is the first time to our knowledge that it has been demonstrated in improvised gesture.

References

- Christensen, P., Fusaroli, R., & Tylén, K. (2016). Environmental constraints shaping constituent order in emerging communication systems: Structural iconicity, interactive alignment and conventionalization. *Cognition*, 146, 67–80.
- Dryer, M. S. (2013). Order of subject, object and verb. In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Gleitman, L. R., January, D., Nappa, R., & Trueswell, J. C. (2007). On the give and take between event apprehension and utterance formulation. *Journal of Memory and Language*, 57(4), 544–569.
- Goldin-Meadow, S., So, W. C., Ozyürek, A., & Mylander, C. (2008). The natural order of events: how speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences of the United States of America*, 105(27), 9163–8.
- Meir, I., Aronoff, M., Börstell, C., Hwang, S.-O., Ilkbasaran, D., Kastner, I., Lepic, R., Lifshitz Ben-Basat, A., Padden, C., & Sandler, W. (2017). The effect of being human and the basis of grammatical word order: Insights from novel communication systems and young sign languages. *Cognition*, 158, 189–207.
- Schouwstra, M., & Swart, H. de. (2014). The semantic origins of word order. *Cognition*, 131(3), 431–6.

PRESSURE FOR TRUST-BASED EFFICIENCY SHAPED THE EVOLUTION OF LANGUAGE

Chris Knight

chris.knight@live.com Department of Anthropology, University College London, London, United Kingdom

For human spoken language to emerge, selection pressures must have acted upon a primate-style precursor that had evolved subjet to the constraints of Darwinian signal evolution theory. Science is a search for conceptual unification. Consistently applied, a single new Darwinian pressure – selection for trust-based efficiency – suffices to explain the core features which distinguish speech from nonhuman primate vocal communication.

1. Introduction

This paper attempts to combine what we know about the evolution of communication in general with the emergence of language in particular, bringing into alignment two apparently irreconcilable theoretical paradigms – Darwinian signal evolution theory on the one hand, historical linguistics including grammaticalization theory on the other.

2. Why apes don't speak

When communication is efficient, detailed information is conveyed accurately and at speed. From a human standpoint, the grunts, calls and screams of primates seem low in conceptual content and strikingly repetitive, as if messages had to overcome resistance from listeners on guard against being tricked. Resistance thwarts efficiency by forcing vocalisers to amplify and repeat. Fast and efficient communication presupposes minimal resistance from listeners who trust that communicative intentions are honest and trickery therefore unlikely.

It used to be thought that chimpanzee vocalizations were limited to involuntary grunts and screams expressing physical and emotional states (e.g. Goodall, 1986: 114-145). But recent studies in the wild have recorded food calls directed at specific individuals, indicating a cooperative intent (Schel et al., 2013a). Vocal alarms may also be volitional and intentional. In one experiment,

two chimps were surprised by a model snake which had been deliberately placed along their path. The one who first noticed began 'hooing' while looking back and forth between his partner and the snake, stopping once confident that the other had seen the object (Schel et al., 2013b). Here, there are no grounds for listener mistrust or suspicion since the referent can be immediately checked out. It would be surprising if the vocalizer 'hooed' to its companion in a cooperative way about an imaginary snake. It would be still more astonishing if the listener heard the 'hoo', saw no snake – yet still played along with the patent fiction.

If great ape vocal communication can be intentional, where exactly is the language frontier which these animals seem unable to cross? Primatologists (e.g. Bergman et al., 2016; Seyfarth and Cheney, 2014, 2016; Vilain et al., 2011) often deny any Rubicon, viewing human capacities for vocal communication as modified extensions of those of our primate cousins or ancestors. In the opposite camp, most theoretical linguists (e.g. Bickerton, 1990; Berwick and Chomsky, 2016) point out that, for all their abilities, our ape relatives have not reached even the first rung of the ladder leading to syntactically complex speech.

3. From a standpoint in signal evolution theory, language should not exist

The faculty of language considered broadly (Hauser et al., 2002) is not one module but an assemblage of previously unrelated parts. The necessary auditory capacities, vocal organs and moveable components such as lips and tongue certainly evolved as adaptations to non-linguistic functions, becoming fine-tuned over evolutionary time to serve radically new ends. It might seem tempting to conceptualize the pressures responsible for this process of exaptation and convergence as 'selection pressures for speech'. But nothing in Darwinian theory permits us to envisage adaptation to achieve some future goal. Adaptation is always to the present, compelling us to envisage selection pressures prior to speech. Instead of assuming linguistic concepts from the outset, therefore, we need to envisage a fundamental pre-linguistic principle which, consistently applied, might have led in a direction toward speech. Since we cannot yet appeal to language, we have no choice but to rely on concepts appropriate within evolutionary biology.

Modern Darwinian signal evolution theory can be traced to the early work of Dawkins and Krebs (1978), who distinguish two opposite selection pressures which determine how signals evolve. One runs counter to efficiency in favouring loud, repetitive, extravagant displays; the other favours speed and efficiency. Normally, conflicts of interest prompt receivers to resist incoming signals, forcing producers to amplify and repeat in order to get through. Although primates in some contexts may engage in so-called 'cheap talk' (Silk et al., 1999), pressure for reliability means that they never get close to the fast, cheap, super-efficient extreme represented by human speech.

The costs to an animal of producing a signal may be divided into two parts – the 'efficacy cost' or investment needed to ensure clear transmission, and the 'strategic cost' needed to demonstrate signal reliability (Maynard Smith and Harper, 2003). To arrive at the fast and efficient extreme represented by speech, the second component needed to be reduced to zero, freeing speakers from all except efficacy costs. This could only happen where listeners were sufficiently trusting, *requiring no demonstration of signal reliability at all*.

Bonobos, chimpanzees and many other primates seem to have capacities for symbolic communication which they don't use in the wild (Fouts 1997; Savage-Rumbaugh 1998). Early hominins must surely have possessed at least equal capacities, yet we find no compelling archaeological evidence for art or symbolic usage until strikingly late, around the time when modern *Homo sapiens* begins to emerge (Henshilwood et al., 2002, 2009; Henshilwood and Dubreuil, 2009; Watts, 1999, 2009, 2014; Watts et al., 2016). This returns us to our opening discussion. There must have been some powerful factor blocking the emergence of either words or grammar throughout the greater part of hominin evolution.

Grammaticalization theory (Meillet, 1903; Heine et al., 1991; Deutscher 2005) has clarified how grammatical structures emerge and systematically change over historical time. Yet despite progress in applying such insights to evolutionary questions (Heine and Kuteva, 2007, 2012; Hurford, 2012) it remains unclear what social, cognitive or neural developments initially enabled grammaticalization to get under way, or what prevented it from happening previously.

Signal evolution theory reminds us of the role which must initially have been played by trust in enabling the first faltering steps toward words and grammar. Without trust in communicative intentions, not even those first steps could have been taken. Where listeners are suspicious from the outset, each successive signal must inspire confidence by demonstrating in some way that it is reliable (Zahavi and Zahavi, 1997), a requirement which slows things down, prompts repetition and impedes efficiency. Repetitive ape hoots, screams and other costly signals are just not the kind of elements that can be reduced, combined, or recursively structured in the manner that grammaticalization requires.

4. Design features of speech

Since the work of Hockett (1960), linguists have catalogued an impressive list of features to distinguish speech from animal systems of communication. Spoken

language is (1) symbolic, (2) metaphorical in principle, (3) grammatically complex, (4) capable of displaced reference and (5) digital in format. Instead of seeking to explain each feature by a separate theory, it would be better if we could explain them all on the basis of a single parsimonious theory.

4.1 Symbolism. Although wild-living apes do not use symbols, this is not because they lack the capacity. Classic studies of captive apes have shown that these intelligent animals have no difficulty in learning and using arbitrary symbols (Fouts, 1997; Savage-Rumbaugh et al., 1998). So why don't they use them in the wild? To explain this, we need to investigate social relations – in particular, primate expectations of honesty and trust.

The philosopher Eco (1976: 7) famously defined a symbol as anything that 'can be used in order to lie'. Yet from a standpoint in primate communication, the problem is that words are *already* 'lies', regardless of their subsequent use. To a wild-living chimpanzee, every vocalization of the kind made in speech would be perceived as a potentially deceptive fake. This is theoretically challenging because, although primate deception can work, it has to be sufficiently rare for victims to be tricked into expecting honesty. A situation in which deceptions had become normal would be hard to explain.

Symbols are often defined as arbitrary social conventions (Peirce, 1940). But this fails to capture the element of falsehood involved. Sperber's definition is more helpful here. "That's symbolic" Why? Because it is false' (Sperber, 1975: 4). A symbol is a falsehood given currency by social convention. The necessary social element, a new kind of trust, has been linked by some scholars to the transition from primate-style dominance politics to the egalitarianism of extant hunter-gatherers (Erdal and Whiten, 1994, 1996; Whiten and Erdal, 2012; Whiten, 1999; Lewis, 2009, 2014).

Before the evolution of language could begin, if these insights are correct, individuals had to start accepting patent falsehoods on trust (Knight, 2008, 2009, 2014; Knight and Lewis, 2014, 2017). We might imagine a hominin 'hooing' to its companion when no snake was visible. If instead of dismissing the false alarm, the listener showed a cooperative interest in imaginary snakes, the way would be open for the multiple complexities of symbolic language to evolve.

4.2 *Metaphor.* The creative use of language depends crucially on the ability to deploy metaphor (Deutscher, 2005; Smith and Hoefler, 2014). Beyond this, metaphor is central to the way humans think (Lakoff and Johnson, 1987). 'Conceptualising one thing while meaning another' is one way to formulate the underlying principle, but metaphorical usage boils down to falsehood: 'Generally it is only when a sentence is taken to be false that we accept it as a

metaphor and start to hunt out the hidden implication' (Davidson, 1979: 40). This hidden implication is a chosen aspect of reality which for joint purposes we feel able to take as truth. When a metaphor becomes common currency, we forget its original incongruity, which was the quality necessary to provoke thought. In becoming conventionally accepted, previously striking metaphors (such as 'bottle neck') become interpreted literally – that is, as settled features of the linguistic code ('bottleneck'). The cyclical logic through which metaphors arise, fade, die and are replaced by novel metaphors accounts quite generally for the creativity of language and its restless unfolding over historical time (Deutscher, 2005).

We might suppose that metaphor requires no special interpretive abilities, just a willingness to accept patent falsehood on trust. But we must not underestimate the cognitive challenges here. For metaphor to work, speakers and listeners must put themselves in each other's shoes, attempting to guess at one another's communicative intentions using contextual cues, memory and imagination (Grice, 1969; Sperber and Wilson, 1986). The collaboration required for the development of such 'intersubjectivity' (Tomasello and Rakoczy, 2003) or 'deep social mind' (Whiten, 1999) is clearly a social and political factor, and it is this which most decisively sets humans in a world apart from apes.

4.3 Grammar. Evolutionary theorists once assumed the need for two separate theories, one to explain the emergence of words in isolation and another to explain how grammar evolved at a later stage (e.g. Bickerton, 1990). The requirement for two stages seemed unavoidable in the days when the dominant theoretical paradigm placed words in a quite separate category from grammatical rules (e.g. Pinker, 1999). Nowadays, most evolutionary linguists embrace some version of construction grammar, which treats words and rules alike as constructions acquired by children simultaneously and in essentially the same way (Tomasello, 2006; Goldberg, 2006). So two separate evolutionary stages are no longer required. Once humans were using metaphor in the broadest sense – once they could *say one thing while meaning another* (Sperber and Wilson, 1987) – words and rules would simultaneously have emerged (Smith and Hoefler, 2014).

4.4 Displaced reference. The ability to refer to events or objects outside the here and now is often conceptualized as a useful additional feature of language. But I prefer to invoke Austin (1978) and Searle (1969), who make the more profound point that *any* speech act is a move internal to a shared virtual (socially constructed). If that is accepted, *all* linguistic reference is of necessity displaced.

Even to point at something is to take action in the world *as jointly construed*, so displacement is involved here, too.

There is no need for a separate theory to explain this. Monkeys and apes inhabit reality, not virtual reality, and so their signalling naturally reflects this. Every primate gesture or vocalization is designed to alter reality as perceived from the standpoint of the signaller. But to inhabit a socially constructed landscape of 'institutional facts' (Searle, 1996) is to be happy with shared fictions, trusting that they are not exploitative tricks. So whatever it was that enabled hominins to accept one another's fictions would simultaneously have enabled displaced reference.

4.5 Digital format. A digital switch is an all-or-nothing choice. Saussure (1983 [1915]) famously observed that in language, there are no positive terms, only categorical differences between them. Jakobson took this insight a step further with his celebrated theory of distinctive features, arguing that the speech sounds of the world's languages reduce to a limited set of binary oppositions, each switching between one phoneme and another and, potentially, between one selected meaning and another (Jakobson et al., 1951). It is worth mentioning that although the vocalizations of some species (e.g. songbirds and cetaceans) display digital structure on the level of sound, digital format does not extend to semantics as well. If a whale or songbird were to replace one note in a sequence by another, this would not systematically modify the meaning of the entire song. With spoken language, by contrast, replacing a voiced by an unvoiced in a sentence might well transform the meaning of the entire utterance.

An ape, on hearing a vocalization, homes in on the multiple acoustic gradations of each incoming sound as it arrives. By contrast, as a guide to meaning, the human speaker-listener just wants to know whether this or that distinct vowel or consonant is intended. A useful way to conceptualize the issues at stake is to contrast speaking with singing. When we sing, our audience will naturally expect the whole performance, listening to the melody for as long as it takes. Resorting to shorthands just won't do. Abbreviation is unthinkable because our focus of interest is not the singer's intentions but their actual performance – the impact of the melody and the precise acoustic quality of those sounds. No performance will satisfy an audience if, on grounds of speed and efficiency, the singer were to compress the output into a sequence of rapid-fire digits.

Although music combines categorically distinct notes and is in that sense a discrete-combinatorial system, where meanings are concerned we remain in a graded, analog realm. Song may have language-like aspects and speech may be melodic (see chapters in Bannan, Ed., 2012; Wallin et al., 2000), yet it is

pressure to transmit conceptual information at speed which drives speech to diverge so radically from music and song. As the extreme limit of efficiency is approached, shorthand abbreviations become reduced to zero-cost digital contrasts.

A digital computer is one which moves by sudden jumps or clicks from one quite definite state to another. Turing (1950) clarifies that, strictly speaking, there are no such machines. Everything in real life moves continuously. Even with electric lighting, explains Turing, 'it is a convenient fiction that each switch must be definitely on or definitely off. There must be intermediate positions, but for most purposes we can forget about them'. The fact that digits appear only when we forget about intermediate states – ignoring irrelevant gradations in sensory experience – tells us that they are not physically real distinctions but hallucinations or abstractions.

Apes are not willing to ignore subtle gradations. The key fact for any evolutionary account is that a nonhuman primate, on hearing conspecifics' graded calls, will not interpret them as abstract digits. Far from ignoring intermediate states, the animal will home in on those subtle gradations which alone can provide information about the quality and reliability of what it hears. Turing's insights help us to appreciate that digital format becomes possible only when listeners feel able to overlook irrelevant performance gradations, vesting their trust, yet again, solely in communicative intentions.

5. Conclusion

Before fictions became acceptable currency, the design features of language reviewed here were barred by the constraints inherent in Darwinian signal evolution theory. Where each signal must demonstrate a strategic cost, not one of language's special features can evolve. From a standpoint in the evolution of animal communication, language therefore seems to be theoretically impossible (Knight, 2016). If, despite this, it did evolve, it is not because Darwinian theory is wrong. Rather, it is that *virtual* action involves no material costs, falling beyond the remit of this whole body of theory. For reasons that I have not explored here (see Knight, 2008, Knight and Lewis, 2014), humans in speech communities vested in one another something approaching infinite trust. This was not because they were unconditionally honest, but simply because all communication was now internal to a shared virtual world.

There is an intimate connection between the intensity of cooperation across a network and levels of honesty and trust. Darwinian theorists (e.g. Nowak, 2006) have developed sophisticated abstract models to explain how cooperation between natural organisms evolves, setting aside complexities such as gender

and politics for the sake of simplicity. When dealing with the evolution of human hunter-gatherer egalitarianism, however, such complexities cannot be ignored. It is beyond the scope of this paper to detail how interpersonal trust would have been raised by limiting within-group sexual competition and conflict, or how evolving human females enhanced their fitness by cementing childcare alliances in ways which also strengthened gender egalitarianism (Knight and Lewis, 2014, 2017). An ambitious intellectual school (Hrdy, 2009; Burkart et al., 2009; Power and Watts, 1996; Power, 2009; Opie and Power, 2009) has recently been applying basic Darwinian theory to the specific case of hominin females with large-brained babies needing to trust one another and cooperate closely in order to nurture their unusually demanding offspring and also to secure provisioning support from males. Here, my aim has been limited to showing how the constraints obstructing language's emergence would have been lifted given sufficient trust. Once there was no need to demonstrate reliability, novel pressures for speed and efficiency would have come into play, leading to the unprecedented design features of speech.

Tolerance of patent falsehood takes listener trust well beyond anything which chimpanzees are prepared to extend to one another. Wild-living chimps have been shown to accept volitional, intentional vocal alarms where their veracity can be simultaneously confirmed. If only they were to take one step beyond that – to the point of welcoming false alarms – they would be crossing the Rubicon into a virtual world.

Science is a search for conceptual unification. The revolutionary new signal evaluation criterion permitting language to emerge was simply this: signals were deemed acceptable even when known to be false. This single step, taken habitually by socially intelligent modern humans, enabled the extraordinary complexities of language to pour out in a cascade.

References

- Austin, J. L. (1978 [1955]). *How to do things with words*. Oxford: Oxford University Press.
- Bannan, N. (Ed.), 2012. *Music, language, and human evolution*. Oxford: Oxford University Press.
- Bergman, T. J., Beehner, J. C., Cheney, D. L., & Seyfarth, R. M. (2003). Hierarchical classification by rank and kinship in baboons. *Science*, *302*, 1234–1236.
- Berwick, R. C. & Chomsky, N. (2016), *Why only us?* Cambridge, MA: MIT Press.
- Bickerton, D. (1990). *Language and species*. Chicago: University of Chicago Press.

- Davidson, R. D. (1979). What metaphors mean. In S. Sacks (Ed.), On metaphor (pp. 29–45). Chicago: University of Chicago Press.
- Dawkins, R. & Krebs, J. R. (1978). Animal signals: Information or manipulation? In J. R. Krebs & N. B. Davies (Eds.), *Behavioural ecology*. *An evolutionary approach* (pp. 262-309). Oxford: Blackwell Scientific Publications.
- Deutscher, G. (2005). *The unfolding of language: The evolution of mankind's greatest invention*. London: Random House.
- Eco, U. (1976). A theory of semiotics. Bloomington, IN: Indiana University Press.
- Erdal, D. & Whiten, A. (1994). On human egalitarianism: An evolutionary product of Machiavellian status escalation? *Current Anthropology* 35(2): 175-183.
- Erdal, D. & Whiten, A. (1996). Egalitarianism and Machiavellian intelligence in human evolution. In P. Mellars & K. Gibson (Eds.), *Modelling the Early Human Mind* (pp. 139-150). Cambridge: McDonald Institute Monographs.
- Fouts, R. S. (1997). *Next of kin: My conversations with chimpanzees.* Harper Collins, New York.
- Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
- Goodall, J. (1986). *The chimpanzees of Gombe. Patterns of behavior*. Cambridge, MA and London: Belknap Press of Harvard University Press.
- Grice, H. P. (1969). Utterer's meanings and intentions. *Philosophical Review* 78, 147-77.
- Hauser, M. D. Chomsky, A. N. & Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science*, *298*, 1569–1579.
- Heine, B., Claudi, U. and Hünnemeyer. F. (1991). *Grammaticalization. A conceptual framework*. Chicago and London: University of Chicago Press.
- Henshilwood, C. S., d'Errico, F., Yates, R., Jacobs, Z., Tribolo, C., Duller, G. A. T., Mercier, N., Sealy, J. C., Valladas, H., Watts, I. & Wintle, A. G. (2002). Emergence of modern human behavior: Middle Stone Age engravings from South Africa. *Science 295:* 1278-1280.
- Henshilwood, C. S. & Dubreuil, B. (2009). Reading the artifacts: gleaning language skills from the Middle Stone Age in southern Africa. In R. Botha & C. Knight (Eds.), *The Cradle of Language* (pp. 41-61). Oxford: Oxford University Press.
- Henshilwood, C., d'Errico, F., & Watts, I., (2009). Engraved ochres from the Middle Stone Age levels at Blombos Cave, South Africa. *Journal of Human Evolution* 57(1): 27-47.
- Hockett, C. (1960). The origin of speech. Scientific American 203(3), 88-96.
- Hrdy, S. B. (2009). Mothers and others. The evolutionary origins of mutual understanding. London & Cambridge, MA: Belknap Press of Harvard University Press.
- Hurford, J. (2012). The origins of grammar. Oxford: Oxford University Press.

- Jakobson, R., C. Gunnar, M. Fant & M. Halle, (1951). Preliminaries to speech analysis: The distinctive features and their correlates. Cambridge, MA: MIT Press.
- Knight, C. (2008). 'Honest fakes' and language origins. Journal of Consciousness Studies, 15 (10–11), 236–48.
- Knight, C. (2009). Language, ochre, and the rule of law. In R. Botha & C. Knight (Eds), *The cradle of language* (pp. 281-303). Oxford: Oxford University Press.
- Knight, C. (2016). Puzzles and mysteries in the origins of language. *Language* and Communication, vol. 50, pp. 12-21.
- Knight, C. & Lewis, J. (2014). Vocal deception, laughter, and the linguistic significance of reverse dominance. In D. Dor, C. Knight & J. Lewis (Eds.), *The social origins of language* (pp. 297-314). Oxford: Oxford University Press.
- Knight, C. & Lewis, J. (2017). Wild voices: Mimicry, reversal, metaphor, and the emergence of language. *Current Anthropology*, *58*(4): 435-453.
- Lakoff, G. & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.
- Lewis, J. (2009). As well as words: Congo Pygmy hunting, mimicry, and play. In R. Botha & C. Knight (Eds.), *The Cradle of Language* (pp. 236-256). Oxford: Oxford University Press.
- Lewis, J. (2014). BaYaka Pygmy multi-modal and mimetic communication traditions. In D. Dor, C. Knight & J. Lewis (Eds.), *The Social Origins of Language* (pp. 77-91). Oxford: Oxford University Press.
- Maynard Smith, J. & Harper, D. (2003). *Animal Signals*. Oxford: Oxford University Press.
- Nowak, M. (2006). Five rules for the evolution of cooperation. *Science*, *314* (5805): 1560-1563.
- Opie, K. & Power, C. (2009). Grandmothering and female coalitions. A basis for matrilineal priority? In N. J. Allen, H. Callan, R. Dunbar & W. James (Eds.), *Early human kinship* (pp. 168-186). Oxford: Blackwell.
- Peirce, C. S. (1940). Logic as semiotic: the theory of signs. In J. Buchler (Ed.), *The philosophical writings of Peirce* (pp. 98-119). New York: Dover.
- Pinker, S. (1999). Words and Rules. The ingredients of language. London: Weidenfeld and Nicolson.
- Power, C. (2009). Sexual selection models for the emergence of symbolic communication: Why they should be reversed. In R. Botha & C. Knight (Eds.), *The cradle of language* (pp. 257-280). Oxford: Oxford University Press.
- Power, C., & Watts, I. (1999). First gender, wrong sex, in H. L. Moore, T. Sanders, & B. Kaare (Eds.), *Those who play with fire: Gender, fertility and transformation in East and Southern Africa* (pp. 101-132). London & New Brunswick, N. J.: Athlone Press.

- Saussure, F. de (1983 [1915]). *Course in general linguistics*. Trans. R. Harris. London: Duckworth.
- Savage-Rumbaugh, E. S. & McDonald, K. (1988). Deception and social Manipulation in Symbol-using Apes. In R. W. Byrne & A. Whiten (Eds.), *Machiavellian intelligence* (pp. 224-237). Oxford: Clarendon Press,
- Schel, A. M., Machanda, Z., Townsend, S. W., Zuberbühler, K. & Slocombe, K E. (2013a). Chimpanzee food calls are directed at specific individuals. In *Animal Behaviour*, 86(5), 955-965.
- Schel, A. M., Machanda, Z., Townsend, S. W., Zuberbühler, K. & Slocombe, K.E. (2013b). Chimpanzee alarm call production meets key criteria for intentionality. *PLOS ONE* 8(10), 1-11
- Searle, J. R. (1969). Speech acts. An essay in the philosophy of language. Cambridge: Cambridge University Press.
- Searle, J. R. (1996). The construction of social reality. London: Penguin.
- Silk, J. B., Kaldor, E. & Boyd, R. (1999). Cheap talk when interests conflict. *Animal Behaviour 59:* 423-432.
- Smith, A. D. M. & Hoefler S. H., (2016). From metaphor to symbols and grammar: the cumulative cultural evolution of language. In C. Power, M. Finnegan & H. Callan (Eds.), *Human origins: Contributions from social* anthropology (pp. 153–179). New York: Berghahn.
- Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge: Cambridge University Press.
- Sperber, D. (1975). *Rethinking symbolism*. Cambridge: Cambridge University Press.
- Sperber, D., & Wilson, D. (1987). *Relevance: Communication and cognition*. Oxford: Blackwell.
- Seyfarth, R. M., & Cheney, D. L. (2014). The evolution of language from social cognition. *Current Opinion in Neurobiology*, 28, 5–9.
- Seyfarth, R. M., & Cheney, D. L. (2016). Precursors to language: Social cognition and pragmatic inference in primates. *Psychonomic Bulletin & Review*, 24(1), 79–84.
- Tomasello, M. (2006). Acquiring linguistic constructions. In D. Kuhn & Siegler, R. (Eds.), *Handbook of child psychology* (pp. 255-98). New York: Wiley.
- Tomasello, M., & Rakoczy, H. (2003). What makes human cognition unique? *Mind and Language 18*:121–147.
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59, 433-460.
- Vilain, A., Schwarts, J.-L. Abry, C. & Vauclair, J. (Eds.), (2011). Primate communication and human language. Vocalization, gestures, imitation and deixis in humans and non-humans. Amsterdam: John Benjamin.
- Wallin, N. L., Merker, B. & S. Brown (Eds.), *The Origins of Music*. Cambridge, MA: MIT Press.

- Watts, I. (1999). The origin of symbolic culture. In R. Dunbar, C. Knight & C. Power (eds), *The Evolution of Culture* (pp. 113-46). Edinburgh: Edinburgh University Press.
- Watts, I. (2009). Red ochre, body painting, and language: Interpreting the Blombos ochre. In R. Botha and C. Knight (Eds.), *The cradle of language* (pp. 62-92). Oxford: Oxford University Press.
- Watts, I. (2014). The red thread: pigment use and the evolution of collective ritual. In D. Dor, C. Knight and J. Lewis (Eds), *The social origins of language* (pp. 208-227). Oxford: Oxford University Press.
- Watts, I., Chazan, M. & Wilkins, J. (2016). Early evidence for brilliant ritualized display: Specularite use in the Northern Cape (South Africa) between ~500 and ~300 Ka. *Current Anthropology*, 57(3).
- Whiten, A. (1999). The evolution of deep social mind in humans. In M. Corballis & S. E. G. Lea (Eds,). *The descent of mind: psychological perspectives on hominid evolution* (pp. 173–193). Oxford: Oxford University Press.
- Whiten, A., & Erdal, D.. (2012). The human socio-cognitive niche and its evolutionary origins. *Philosophical Transactions of the Royal Society B: Biological Sciences 367:* 2119–2129.
- Zahavi, A. and A. Zahavi (1997). *The handicap principle. A missing piece in Darwin's puzzle*. New York & Oxford: Oxford University Press.

DO CHILDREN USE "FINE" OSTENSIVE COMMUNICATION?

HARUMI KOBAYASHI^{*1}, and TETSUYA YASUDA¹

*Corresponding Author: h-koba@mail.dendai.ac.jp ¹Tokyo Denki University, Saitama, Japan

Ostensive communication is a type of communication in which the speaker presents certain information with communicative intention, i.e., with evidence of intention to convey information to the receiver (Wilson & Sperber, 2012). Ostensive communication is claimed to be unique to human communication (Senju & Csibra, 2008; Tomasello, 2008) and it is possibly the foundation of human language (Scott-Phillips, 2015). Ostensive communication about specific aspects of the environment is essential. For example, we need to correctly specify and convey information about parts, including object parts, of the environment in many situations, such as "The handle of the knife is fragile, so it needs fixing" or "The dog's tail may be injured, so it needs care." We wish to claim that humans needed more "fine" ostensive communication skills to explicitly convey information on object parts, and this is one reason why language evolved.

If fine ostensive communication is important for the evolution of language, the course of the development of these skills in children may provide precious information about the mechanism of ostensive communication. We examined 4and 6-year-olds' ability to recognize and control ostensive cues using the "teaching part names paradigm" that we originally developed in this study. Teaching seems to be an ultimate form of ostensive communication because, in addition to presenting the clear intention of conveying information to the receiver, the speaker must consider the receiver's viewpoint, knowledge, and skills, and appropriately control her own teaching behavior.

Eleven 4-year-old and 12 6-year-old preschoolers participated in the experiment. Two types of pointing gestures, namely, touch-pointing and 7 cm distance pointing, were used in the experimental task. In this task, the experimenter introduced a lion puppet as a language learner and asked the child to teach certain word meanings to the puppet. Then the experimenter showed the child an experimental puppet (e.g., a bear) that holds an object (e.g., a patrol car). The experimenter then pointed with either touching or without touching (with 7 cm distance to the object part) one of the tires of the patrol car with her

finger, and said in Japanese, "This is *agu* (nonsense syllable) in a foreign language." Thus, the critical object part was embedded in an object that was held by a puppet. Then she asked the child to teach the nonsense syllable to the "learner puppet" (production trial). The child was also asked to respond to "What is *agu*?" either using language or gestures (comprehension trial). The child's teaching behavior (pointing with touching or close to the object part/pointing without touching the object part/showing to the puppet/other) and responses were analyzed.

In the gesture production trial, a 2 (Age: 4 and 6 year-olds) \times 2 (Teaching: 7-cm pointing and touch-pointing) mixed two-way ANOVA was performed, with each type of gesture when the children taught part names (touch-pointing, showing) as dependent measures. Other gestures were rare, and therefore, they were excluded from the present analysis. There was a significant main effect of age and a significant interaction of age \times teaching. The simple main effect of the age \times teaching interaction revealed that 6-year-olds pointed to the object more frequently than the 4-year-olds did in the 7-cm pointing teaching. In the comprehension trial, 6-year-olds successfully guessed more part names when the experimenter touch-pointed at the object part than when she engaged in 7-cm pointing.

The present findings showed that 6-year-olds better understand and control ostensive cues in teaching word meanings than 4-year-olds do. The older children seemed to implicitly know that different pointing gestures are appropriate for teaching part and whole labels. However, even 6-year-olds who understood the meaning of touch-pointing did not touch-point at the object part producing the part name if the object part was presented in the embedded situation. This result suggests that "fine" ostensive communication, including the appropriate use of showing, is in the developmental state in 4-year-olds. The study also seems to suggest that ostensive cue control precedes language control.

References

- Scott-Phillips, T. (2015). Speaking our minds: Why human communication is different, and how language evolved to make it special. UK: Palgrave MacMillan.
- Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. Current Biology, 18(9), 668-671.
- Tomasello, M. (2008). Origins of human communication. Cambridge, MA: Harvard University Press.
- Wilson, D., & Sperber, D. (2012). Meaning and relevance. Cambridge: Cambridge University Press.

Acknowledgements

This work was supported by MEXT/JSPS KAKENHI Grant Number JP17H06382 in #4903 (Evolinguistics), JP 16K04318, and JP 17K04949.
RECONTEXTUALIZATION: THE DYNAMICS OF LANGUAGE BEHAVIOR AND CHANGE

ANNELIESE KUHLE*1 and FERDINAND VON MENGDEN1

*Corresponding Author: anneliese.kuhle@fu-berlin.de ¹Institute for English Language and Literature, Freie Universität Berlin, Berlin, Germany

Lévi-Strauss (1966 [1962]) famously characterized human behavior as acts of *bricolage*. In later years, following Lévi-Strauss's research, this notion has become known as *recontextualization*—the label we adhere to in our own presentation. This notion is defined as the ability of subjects to constantly re-use available materials based on prior experience in a context-dependent manner. For instance, a block of wood can be re-used by human individuals as a 'stool' to sit on, a 'table' to sit at, or just as well, depending on the context and intention of the individual, as a support item for other material. If such behavior is placed in a collective, social environment, it provides a crucial platform for the incidental emergence of social conventions—varying not only across groups but also among individuals within groups. One such set of conventions, we believe, is human language and its structures.

In our talk, we would like to argue that recontextualization already occurs in pre-linguistic behaviors of nonhuman primates, and further that this finding offers new insights for the discussion on language evolution and change.

In a first step, we present recontextualization as a phenomenon characterizing primitive tool uses of pre-linguistic animals such as great apes (Kuhle, 2014, p. 3). Primatologists, who investigate such behavior both experimentally and under natural conditions, identify instances of recontextualization based on *means-ends dissociation*, i.e., the re-use of old means from prior experience in novel contexts of behavior—which thereby unintentionally leads to new form-function pairings (Boesch, 2013; Call, 2013). Such primitive tool uses are *context-dependent* both with regard to their "invention" and subsequent social transmission within the group. They vary across groups to such a degree that subjects from different communities can be

identified by their tool conducts rather than their physical appearance (McGrew, 1998; Whiten et al., 1999).

In a second step, we argue that the study of language within linguistics can benefit from the notion of recontextualization. We consider instances of grammatical change (e.g., the various uses of English *let's/lets* and the emergence of a new imperative marker) in order to show how acts of recontextualization in linguistic usage explain how different variants and ultimately new linguistic patterns emerge. Such new patterns can also be characterized in terms of means-ends dissociation with regard to linguistic form-function pairings. Speakers thus create new patterns without deviating from patterns that already exist in their linguistic experience.

We agree with other usage-based approaches that language is a complex adaptive system with dynamic and emergent properties (see, e.g., Heine, 2002; Beckner et al., 2009; Winters, 2010; Steels, 2011). However, we differ from some of these approaches by introducing a more radical notion of contextdependence and 'emergence' of linguistic behavior. Our claim is that the linguistic sign is inherently negotiable, underspecified and subject to interpretation and that variation and ambiguity are inherent properties of the linguistic sign. We therefore consider any approach which conceptualizes 'change' as a transition from one 'synchronic state' (stage) to another as problematic. Natural, real-time dynamics of linguistic activity do not attest to and do not require these assumptions, but they do attest to behavioral strategies that can be described as 'recontextualization'.

Our approach has obvious consequences for the question of language evolution. complex linguistic structures emerge via acts If of recontextualization, and if comparable acts of recontextualization are observed in non-linguistic behavior both among humans (bricolage) and among nonhuman primates, then we have identified a behavioral strategy that underlies the transition from non-linguistic forms of communication (or acting in general) to natural languages. We argue that this observation constitutes the common denominator that not only explains linguistic structures, but also connects linguistic behavior with other types of non- or pre-linguistic behavior.

References

Beckner, C., Blythe, R., Bybee, J., Christiansen, M. H., Croft, W., Ellis, N. C., Holland, J., Ke, J., Larsen-Freeman, D., Schoenemann, T. (2009). Language is a complex adaptive system: Position paper. *Language Learning*, 59, 1-26.

- Boesch, C. (2013). Ecology and cognition of tool use in chimpanzees. In C. M. Sanz, J. Call and C. Boesch (Eds.), *Tool use in animals: Cognition and ecology* (pp. 21-47). Cambridge: Cambridge University Press.
- Call, J. (2013). Three ingredients for becoming a creative tool user. In C. M. Sanz, J. Call and C. Boesch (Eds.), *Tool use in animals: Cognition and ecology* (pp. 3-20). Cambridge: Cambridge University Press.
- Heine, B. (2002). On the role of context in grammaticalization. In I. Wischer and G. Diewald (Eds.), *New reflections on grammaticalization* (pp. 83-101). Amsterdam: John Benjamins.
- Kuhle, A. (2014). Language as tool: The analogy to primate cognition. Language & Communication, 34, 1-16.
- Lévi-Strauss, C. (1966 [1962]). *The savage mind (La pensée sauvage)*. The Nature of Human Society Series. London: Weidenfeld & Nicolson.
- McGrew, W. C. (1998). Culture in nonhuman primates? Annual Review of Anthropology, 27, 301-328.
- Steels, L. (2011). Modeling the cultural evolution of language. *Physics of Life Reviews*, 8, 339-356.
- Whiten, A., Goodall, J., McGrew, W. C., Nishida, T., Reynolds, V., Sugiyama, Y., Tutin, C. E. G., Wrangham, R. W., Boesch, C. (1999). Cultures in chimpanzees. *Nature*, 399, 682-685.
- Winters, M. (2010). Introduction: On the emergence of diachronic cognitive linguistics. In M. E. Winters, H. Tissari and K. Allan (Eds.), *Historical cognitive linguistics* (pp. 3-28). Berlin: Mouton de Gruyter.

MAMMALIAN MODELS FOR THE STUDY OF VOCAL LEARNING: A NEW PARADIGM IN BATS

ELLA Z. LATTENKAMP $^{1,2},$ SONJA C. VERNES $^{1,\ast,\dagger},$ and LUTZ WIEGREBE 2,†

*Corresponding author: sonja.vernes@mpi.nl ¹Neurogenetics of Vocal Communication, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands ²AG Wiegrebe, Dept. Biology II, Ludwig-Maximilians University, Munich, Germany [†]Authors contributed equally

1. Animal models for the study of human speech

The complexity of human spoken language in structure and expression is unmatched in the animal kingdom. However, several animal species possess the capacity for vocal production learning (in the following simply referred to as 'vocal learning'), which is a requirement for the acquisition and production of human speech. Vocal learning is defined as the ability to imitate new sounds or modify existing signals from a preexisting repertoire according to auditory input (Boughman & Moss, 2003; Janik & Slater, 1997). Even though on the surface this may seem simple enough, this is a complex trait that requires the functional interplay of precise auditory perception, neuronal integration of the signal (including template memorization and matching), and accurate neuronal motor control of the vocal production systems (Fitch, 2000). In order to gain an indepth understanding of the evolution and the biological basis of human spoken language, in vivo experiments are an essential prerequisite. As these are challenging in humans, animal models are of great importance for the experimental approach to answering questions about the origins of human language and speech acquisition.

1.1. Vocal Learning in vertebrates

Vocal learning as a behavioural trait can be found in different levels of expression within the animal kingdom and should be regarded as a continuum (Petkov &

Jarvis, 2012). Several vertebrate species have shown indications for vocal learning in the past. These include three groups of birds (songbirds, hummingbirds, and parrots) and several mammalian taxa (cetaceans, pinnipeds, elephants, and bats). In the past decades, songbirds, which occupy a position toward the upper end of the vocal learning continuum, have been studied extensively and this research has yielded invaluable insights in the physiological prerequisites, neurobiology, and genetics required for vocal learning (Jarvis et al., 2005; Pfenning et al., 2014; Roberts et al., 2017; Wheatcroft & Qvarnström, 2015). While research on avian model organisms dominates the field, mammalian vocal learning by comparison has largely been neglected in the past.

1.2. Bats as a mammalian model for Vocal Learning

Comparing the available sample sizes, ease of study, and potential different mammalian vocal learners provide, bats present themselves as a useful candidate model system for the study of vocal learning. Bats are gregarious, highly vocal animals with a broad vocal repertoire. In the past, several bats species have shown indications of vocal learning (Knörnschild, 2014; Prat, Taub & Yovel, 2015). However, controlled, repeatable experiments with isolated adult bats have thus far been lacking.

We have developed an automated setup and a conditioning training regime on social vocalisations of isolated bats, which allow for the spectro-temporal tracking of call parameters and their changes over time. This approach for the first time enables the detailed study of vocal parameters and behaviours on an individual level. We were able to demonstrate the volitional use of communication calls out of context and in the absence of any social feedback, hence, capturing usage learning in isolated bats. Furthermore, by changing the reward schedule, we induced temporal and spectral adjustment of call parameters, showing vocal plasticity in bat social calls. These studies of individual bats will provide the basis for the in-depth study of vocal development, and vocal learning in a mammalian model species. Repeatable experiments on isolated individuals will further allow the investigation of common morphological structures among vocal learners, shared auditory processing pathways, and neuronal and gene networks underlying the mammalian vocal learning capacity.

Our experimental approach demonstrates the value of bats as a mammalian model system for vocal learning and will take us a step forward in understanding the evolution of human speech and language acquisition.

Acknowledgements

This work was funded by a Human Frontiers Science Program (HFSP) Research grant (RGP0058/2016) awarded to L.W. and S.C.V. and a Max Planck Research Group Grant awarded to S.C.V.

- Boughman, J.W. & Moss, C.F. (2003). Social sounds: vocal learning and development of mammal and bird calls. In: A. M. Simmons, R. R. Fay, and A. N. Popper (Eds), *Acoustic Communication - Springer Handbook of Auditory Research* (pp. 138–224). Springer-Verlag: New York, Inc.
- Fitch, W.T. (2000). The evolution of speech: a comparative review. *Trends in Cognitive Sciences*, 4, 258–267.
- Janik, V.M. and Slater, P.J.B. (1997). Vocal learning in mammals. *Advances in the Study of Behaviour*, 26, 59–99.
- Jarvis, E.D.; Güntürkün, O.; Bruce, L.; Csillag, A.; Karten, H.; Kuenzel, W.; Medina, L.; Paxinos, G.; Perkel, D.J.; Shimizu, T.; Striedter, G.; Wild, J.M.; Ball, G.F.; Dugas-Ford, J.; Durand, S.E.; Hough, G.E.; Husband, S.; Kubikova, L.; Lee, D.W.; Mello, C. V.; Powers, A.; Siang, C.; Smulders, T. V.; Wada, K.; White, S.A.; Yamamoto, K.; Yu, J.; Reiner, A. and Butler, A.B. (2005). Avian brains and a new understanding of vertebrate brain evolution. *Nature Reviews Neuroscience*, 6, 151–159.
- Knörnschild, M. (2014). Vocal production learning in bats. Current Opinion in Neurobiology, 28, 80–85.
- Petkov, C.I. and Jarvis, E.D. (2012). Birds, primates, and spoken language origins: behavioral phenotypes and neurobiological substrates. *Frontiers in Evolutionary Neuroscience*, 4, 1–24.
- Pfenning, A.R.; Hara, E.; Whitney, O.; Rivas, M. V; Wang, R.; Roulhac, P.L.; Howard, J.T.; Wirthlin, M.; Lovell, P. V; Ganapathy, G.; Mouncastle, J.; Moseley, M.A.; Thompson, J.W.; Soderblom, E.J.; Iriki, A.; Kato, M.; Thomas, M.; Gilbert, P.; Zhang, G.; Bakken, T.; Bongaarts, A.; Bernard, A.; Lein, E.; Mello, C. V; Hartemink, A.J. and Jarvis, E.D. (2014). Convergent transcriptional specializations in the brains of humans and song-learning birds. *Science*, 346, 1–13.
- Prat, Y.; Taub, M. and Yovel, Y. (2015). Vocal learning in a social mammal: demonstrated by isolation and playback experiments in bats. *Science Advances*, 1, 1–5.
- Roberts, T.F.; Hisey, E.; Tanaka, M.; Kearney, M.G.; Chattree, G.; Yang, C.F.; Shah, N.M. and Mooney, R. (2017). Identification of a motor-to-auditory pathway important for vocal learning. *Nature Neuroscience*, 20, 978–986.
- Wheatcroft, D. and Qvarnström, A. (2015). A blueprint for vocal learning: auditory predispositions from brains to genomes. *Biology letters*, 11, 20150155.

SEQUENCING BAT GENOMES TO SHED LIGHT ON LANGUAGE EVOLUTION

KSENIA LAVRICHENKO^{*1}, SONJA C. VERNES^{1,2}

*Corresponding Author: ksenia.lavrichenko@mpi.nl ¹Neurogenetics of Vocal Communication Group, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

²Donders Centre for Cognitive Neuroimaging, Nijmegen, The Netherlands

1. Summary

Language is characterized by a number of biological traits, many of which are unique to humans, however some are shared with a handful of other species. These shared traits present an opportunity for language researchers to use cutting-edge methods on relevant model species to study the biological underpinnings of language and its evolution. One such shared, language-relevant trait is vocal learning; the ability to change or create completely new sounds based on auditory feedback (Janik & Slater, 2000; Petkov & Jarvis, 2012).

Vocal learning birds are the most well studied species, and recent efforts to perform large scale sequencing of bird genomes has led to advances in our understanding of how this trait is encoded in a genome (Pfenning et al., 2014). Similarly it has opened the door to molecular and neurogenetic interrogations in both learning and non-vocal learning species (Haesler et al., 2007; Whitney et al, 2014) that will help demonstrate what is necessary and/or sufficient for the evolution of a vocal learning brain/animal.

To gain a broader view of how genomes evolved to facilitate vocal learning, it would be of great value to address such questions in other, more distantly related vocal learning species. Bats have recently caught the attention of the field due to their potential for vocal learning, sophisticated social structure and diversity (Knörnschild, 2014). Two species in particular have recently shown promise for vocal learning; *Phyllostomus discolor* (Esser, 1994) and *Rousettus*

aegypticaus (Prat et al., 2015). Studying the genomes of these species provides a great opportunity to pinpoint genetic loci and major evolutionary events that may be relevant for the emergence of vocal learning in mammals. Furthermore, this will facilitate in depth molecular and neurogenetic studies to understand the biological encoding of this trait in a mammalian brain.

We will report our work using existing and newly generated sequence data to assemble complete genomes and transcriptomes for both these species and the potential it has to answer wider queries about the encoding and evolution of vocal learning. With such data at hand, a number of questions can be addressed, such as:

- 1. What are the shared features of a vocal learning genome, i.e. what are the sequence evolution patterns within the genetic loci relevant to vocal learning both within bat genomes and between bats and other groups, such as birds?
- 2. Is there evidence of functional evolution in these loci, taken into account the higher than expected sequence variability in bats compared to other species? (Li et al., 2007)
- 3. What are the patterns of gene expression in vocal learning-relevant regions of the brain?
- 4. What are the genes and gene networks that play role in vocal learning in the brains of these bats? (Rodenas-Cuadrado et al., 2015).

Studies addressing these and follow up questions will build a sound foundation for elucidating molecular and genetic bases of vocal learning in bats and consequently in humans, as well as providing hints on how this trait emerges in evolution.

Ethical statement

All tissue was obtained from captive colony animals that were kept in compliance with the principles of laboratory animal care and the regulations of the relevant national laws for animal Protection.

Acknowledgements

This work was funded by a Marie Curie Career Integration Grant (PCIG12-GA-2012-333978), and a Max Planck Research Group Grant both awarded to S.C.V. This work was also funded by a Human Frontiers Science Program (HFSP) Research grant (RGP0058/2016) and is part of a collaborative project involving Prof. Lutz Wiegrebe (LMU, Germany), Prof Michael Yartsev (UC Berkeley, USA) and Prof Uwe Firzlaff (TUM, Germany). With thanks to Prof Erich D.

Jarvis (Rockefeller University, USA), the G10k Vertebrate Genome Laboratory and Olivier Fedrigo (Rockefeller University, USA) for sequencing and assembling the genomes; and to the directors of Bat1k consortium for their contributions to this project.

- Janik, V. and Slater, P. (2000). The different roles of social learning in vocal communication. *Animal Behaviour*, 60(1), pp.1-11.
- Petkov, C. and Jarvis, E. (2012). Birds, primates, and spoken language origins: behavioral phenotypes and neurobiological substrates. *Frontiers in Evolutionary Neuroscience*, 4.
- Pfenning, A., Hara, E., Whitney, O., Rivas, M., Wang, R., Roulhac, P., Howard, J., Wirthlin, M., Lovell, P., Ganapathy, G., Mountcastle, J., Moseley, M., Thompson, J., Soderblom, E., Iriki, A., Kato, M., Gilbert, M., Zhang, G., Bakken, T., Bongaarts, A., Bernard, A., Lein, E., Mello, C., Hartemink, A. and Jarvis, E. (2014). Convergent transcriptional specializations in the brains of humans and song-learning birds. *Science*, 346(6215), pp.1256846-1256846.
- Haesler, S., Rochefort, C., Georgi, B., Licznerski, P., Osten, P. and Scharff, C. (2007). Incomplete and Inaccurate Vocal Imitation after Knockdown of FoxP2 in Songbird Basal Ganglia Nucleus Area X. *PLoS Biology*, 5(12), p.e321.
- Whitney, O., Pfenning, A., Howard, J., Blatti, C., Liu, F., Ward, J., Wang, R., Audet, J., Kellis, M., Mukherjee, S., Sinha, S., Hartemink, A., West, A. and Jarvis, E. (2014). Core and region-enriched networks of behaviorally regulated genes and the singing genome. *Science*, 346(6215), pp.1256780-1256780.
- Knörnschild, M. (2014). Vocal production learning in bats. Current Opinion in Neurobiology, 28, pp.80-85.
- Esser, K. (1994). Audio-vocal learning in a non-human mammal. *NeuroReport*, 5(14), pp.1718-1720.
- Prat, Y., Taub, M. and Yovel, Y. (2015). Vocal learning in a social mammal: Demonstrated by isolation and playback experiments in bats. *Science Advances*, 1(2), pp.e1500019-e1500019.
- Li, G., Wang, J., Rossiter, S., Jones, G. and Zhang, S. (2007). Accelerated FoxP2 Evolution in Echolocating Bats. *PLoS ONE*, 2(9), p.e900.
- Rodenas-Cuadrado, P., Chen, X., Wiegrebe, L., Firzlaff, U. and Vernes, S. (2015). A novel approach identifies the first transcriptome networks in bats: a new genetic model for vocal communication. *BMC Genomics*, 16(1).

THE METHOD OF VALIDATION BY ZENITH

DAVID A. LEAVENS*1

*Corresponding Author: davidl@sussex.ac.uk ¹School of Psychology, University of Sussex, Brighton, United Kingdom

A central question in consideration of the origins of language is to what extent linguistic communication reflects motivational and cognitive factors that evolved uniquely in the human lineage. Theorists have argued that humans have unique motivations for sharing attention and for collaborative action (shared intentionality), and uniquely human perceptions of psychological common ground, which are evident in the social interactions of human children, even before they master speech (e.g., Tomasello, Carpenter, & Liszkowski, 2007). Others have claimed that humans have unique representational capacities that are manifested early in childhood (e.g., Povinelli, Bierschwale, and Čech, 1999). For over 20 years, these claims for human cognitive exceptionalism have relied on differences in response profiles between young humans and substantially older great apes when challenged with tests of their social awareness—age differences are confounded with species classifications (Leavens, Bard, & Hopkins, 2017).

In this talk, I will describe the method of Validation by Zenith; this technique identifies a maximum capability response profile in humans, against which the performances of younger humans and animals can be compared. This method assumes that human adults respond to cognitive challenge with the most sophisticated psychological processes in the animal kingdom. This corrects for a bias that exists in the contemporary literature: performance differences between human children and older apes are interpreted as evidence of cognitive superiority of human children, but these interpretations are not validated against response profiles of human adults. If human adults respond similarly to human children in these tasks, then this validates interpretations of human children's cognitive superiority. On the other hand, if human adults behave similarly to apes, in a range of contemporary cognitive assays, then this refutes assumptions of the cognitive superiority of human children over older apes.

For example, Povinelli et al. (1999) claimed that human children, but not chimpanzees, understood gaze direction as a referential signal. The human children were aged 3 years, whereas the chimpanzees were aged 6 years—twice as old as the human children. The general approach of this study was to challenge participants to use human adult behavioral cues to identify the location of hidden objects from two possible hiding locations. In a critical test condition, the experimenter turned to fixate a point significantly above the baited container, but in the same hemispace. In this condition the chimpanzees outperformed the human children, who performed at chance levels. The authors claimed two things about these results: first, they claimed that they had predicted the result in advance of the study on theoretical grounds and, second, they claimed that the higher performances of the chimpanzees reflected a "low-level" understanding of visual attention in others.

This interpretation was challenged by the later application of the method of Validation by Zenith, developed (albeit not named as such) by Thomas, Murphy, Pitt, Rivers, & Leavens (2008). In Thomas et al., the experimenters made the assumption that human adults would exercise the maximum possible cognitive sophistication in the use of experimenter-given cues. They found that human adults, like the 6-year-old chimpanzees in Povinelli et al. (1999), successfully used the averted gaze to the correct hemisphere to locate hidden food.

I will describe more recent applications of the method of Validation by Zenith, which further support the conclusion that previous reports of cognitive differences between young humans and older apes implicate age differences, not evolutionarily based differences in cognitive skills in these kinds of cueing tasks.

- Leavens, D. A., Bard, K. A., & Hopkins, W. D. (2017). The mismeasure of ape social cognition. *Animal Cognition*. https://doi.org/10.1007/s10071-017-1119-1 (Published online in advance of print.)
- Povinelli, D. J., Bierschwale, D. T., & Čech, C. G. (1999). Comprehension of seeing as a referential act in young children, but not juvenile chimpanzees. *British Journal of Developmental Psychology*, 17, 37–60.
- Thomas, E., Murphy, M., Pitt, R., Rivers, A., & Leavens, D. A. (2008). Understanding of visual attention by adult humans (*Homo sapiens*): A partial replication of Povinelli, Bierschwale, and Čech (1999). *Journal of Comparative Psychology*, 122, 428-436.
- Tomasello, M, Carpenter, M., Liszkowski, U. (2007). A new look at infant pointing. *Child Development*, 78, 705-722.

WHAT 50 MILLION DRAWINGS CAN TELL US ABOUT SHARED MEANING

MOLLY LEWIS*1,2 and GARY LUPYAN1

*Corresponding Author: mollyllewis@gmail.com
¹Department of Psychology, University of Wisconsin-Madison
²Computation Institute, University of Chicago

A foundational assumption of linguistic communication is that conversants have similar underlying concepts (Brennan & Clark, 1996; Wierzbicka, 2012). On this view, the ability of one person to understand another when she says "the tree" depends on the word activating the same concept in both people. One approach to verifying this assumption is to rely on definitions, but this reasoning is circular—how can we be sure the words *in* our definitions are the same? Here, we investigate the assumption of shared linguistic concepts by studying concepts represented in the visual modality—drawings—and examining predictors of their variability. Specifically, we ask whether people who are geographically closer and inhabit a similar linguistic environment produce more similar drawings.

We analyzed a dataset of 50 million drawings (of mostly concrete artifacts such as "tree") from 15 million participants (QuickDraw: quickdraw.withgoogle.com/data). Although all drawings were elicited in English, the participants spanned the globe and, we can assume, represent a wide variety of cultural and linguistic experience. Such drawings only capture a part of meaning people know much more about trees than what they look like—and therefore offer a conservative estimate of diversity.

We quantified similarity of drawings in two ways: (1) Hausdorff Distance (HD), which quantifies image similarity as the minimum Euclidean distance between two sets of points (Huttenlocher, Klanderman, & Rucklidge, 1993); (2) the internal weights (layer FC2) for each of our drawings from a neural net model trained on ImageNet (Deng, 2009), with similarity corresponding to the cosine distance (CD) on weights between pairs of images.

Initial analysis included 1500 image pairs of two categories—"bread" and "tree"—from participants located in 72 countries. We validated our similarity measures using human judgements. We selected 20 pairs from each HD decile for each category (Fig. 1a), and asked participants to rate the similarity of the objects in the drawings using a 1 (almost identical) to 7 (completely different) Likert scale. Each participant (N = 100) rated 50 pairs drawings from the same category.



Figure 1. (a) Sample drawing pairs ("bread" and "tree") used in the human similarity judgement experiment. Each row corresponds to a Hausdorff decile. (b) The prototypical "bread" drawing for each country, calculated as the drawing with the smallest average pairwise distance to other drawings from the same country.

Human judgements of similarity were highly correlated with HD (r = .29, p < .0001) as well as CD (r = 0.20, p < .0001). In a mixed effect model with HD and CD as fixed effects, the two measures were simultaneously predictive of human similarity judgements (HD: $\beta = .35$; t = 12.39; CD: $\beta = -.26$; t = -9.95) and thus appeared to capture different aspects of visual similarity.

With our automated similarity measures validated, we next examined predictors of variability in drawing similarity. We sampled 42,900 pairs of drawings across countries for "bread," "tree" and 15 additional items, and then calculated the HD for each pair. We quantified geographic distance as the distance in meters between the centroid of each pair of countries. We quantified linguistic distance in two ways: (1) vocabulary overlap (ASJP database; Bakker, et al., 2009, Dediu, in press); (2) grammatical similarity based on features values from the WALS typological database (Dediu, in press). The best fitting model revealed an effect of all three distance measures on picture similarity, and pointed to an interaction between grammatical similarity, vocabulary overlap and geographical distance: For languages that differed more in terms of their grammar, countries with greater overlap in vocabulary ($\beta = .002$; t = 3.27) or smaller geographic distance tended to have more similar drawings ($\beta = .002$; t = -2.17).

These data reveal systematic cross-cultural variability in semantics, and suggest that speakers' physical and linguistic proximity may contribute to convergence on shared semantics.

- Bakker, D., Müller, A., Velupillai, V., Wichmann, S., Brown, C. H., Brown, P., Egorov, D., Mailhammer, R., Grant, A., & Holman, E. W. (2009). Adding typology to lexicostatistics: A combined approach to language classification. *Linguistic Typology*, 13(1), 169–181.
- Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(6), 1482.
- Dediu, D. (in press). Making genealogical language classifications available for phylogenetic analysis: Newick trees, unified identifiers, and branch length. *Language Dynamics and Change*.
- Deng, J. (2009). Imagenet: A large-scale hierarchical image database. 248-255.
- Huttenlocher, D. P., Klanderman, G., & Rucklidge, W. (1993). Comparing images using the hausdorff distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *15*, 850–863.
- Wierzbicka, A. (2012). Understanding others requires shared concepts. *Pragmatics & Cognition*, 20(2), 356–379.

WHAT DO ICONICITY JUDGEMENTS REALLY MEAN?

Hannah Little1 and Justin Sulik2

*Corresponding Author: hannah.little@uwe.ac.uk ¹University of the West of England, Bristol.
²Royal Holloway, University of London. justin.sulik@gmail.com

Iconicity is a central topic in language evolution research. It has been studied as a tool for bootstrapping communication systems (Fay et al., 2013), as an emergent property (Tamariz et al., 2017), and as a confound in the emergence of structure (Verhoef et al., 2016). Here, we illustrate potential issues with methods used to measure iconicity in language evolution experiments, both those identified in existing published examples, and present two new case studies. We specifically focus two confounds: world knowledge and modality-specific effects of salience.

Iconicity judgements fall into two broad classes. One method is to have naïve participants **match** a signal with a meaning from an array (Perlman et al., 2015). If a signal is iconic, then participants are more likely to match it with its intended meaning than with other meanings. Another method is to ask participants to **rate** how well the signal resembles its referent (Perry et al., 2015).

Previously discussed confounds in iconicity judgement tasks include language knowledge, and the instructions given to participants providing the iconicity judgements. Occhino et al. (2017) looked at the difference between iconicity judgements given for existing signs in both American Sign Language and German Sign Language by native signers of each language. They found that participants rated signs from their native languages as being significantly more iconic than signs in the other language. Ortega et al. (2017) showed that spontaneous gestures that have overlap with linguistic signs are rated more iconic by native signers than by non-signers. These results show that linguistic knowledge has a significant effect on how participants respond in rating experiments. Perry et al. (2015) explored the effects of different sets of instructions in rating tasks. They found that specific words are judged to be more iconic when participants are asked to rate how likely an alien is to guess the meaning of a word, compared to simply rating resemblance between words and meanings. Importantly, context can also contribute to a word being guessable. For example, the first thing someone says on meeting you being a greeting is guessable regardless of how iconic the signal is.

In our first case study (Sulik, 2018), we present data from a graphical signalling task based on Fay et al. (2010). Iconicity ratings represent how well the

signal resembles the referent. While the results, discussed above, from Occhino et al. (2017) focus on language knowledge, the current study focuses on world knowledge. At first glance, the results from (Sulik, 2018) suggest there is a significant effect of iconicity on success: the more iconic the signal, the more likely participants were to guess its meaning. However, we show that the iconicity ratings represent a fairly superficial predictor of performance, and that a better predictor of success is a measure of the complexity of the relevant world knowledge. Some representations are relatively simple (there is one obvious way to construe something; people share that construal). Other representations are more complex (people don't share intuitions about which of several construals is the most salient). The conclusion here is not that iconicity ratings are uninformative or misguided, but rather that the iconicity of a signal is tightly bound up with the world knowledge motivating that signal. Thus, a statistical relationship between iconicity ratings and performance does not necessarily mean that iconicity explains performance. The take-home message is that language evolution experiments that rely on iconicity ratings should control for the complexity of the relevant meanings, and we illustrate one simple way to control for this.

In our second case study, we discuss a matching task. (Little, Eryılmaz, & Boer, 2017) used a forced-choice matching task to measure iconicity in artificial signals produced during a communication game, where participants were asked to select a signal's meaning from 4 choices. Some signals were more commonly correctly matched with their referents than others, indicating that those signals were more iconic. Here, we present new data on how many times listeners clicked on particular meanings in the playback experiment, regardless of what the correct meaning was. This data is useful to investigate whether some meanings are simply selected more often overall because they lend themselves particularly well to the iconicity afforded by the particular modality being used. While we find that there was no correlation between how often signals were clicked overall and how often they were correctly matched by naïve listeners, some meanings were clicked on substantially more often than others, indicating that participants thought many signals resembled those particular meanings. We interpret this as an effect of modality, cohering well with previous findings. For example, in a forced-choice task, participants were more likely to interpret gestures as verbs than nouns because gesture is an action-orientated modality (Micklos, 2017). An alternative interpretation is that some meanings may be chosen more than others just because that meaning is very salient and attracts participants' attention. Regardless of interpretation, the affordances of the modality used in the experiment can cause some meanings to be chosen more than others. Accordingly, we urge that matchingbased language evolution experiments control for choice frequency.

Overall, we present evidence that extends previously observed knowledge- and modality-based confounds, and highlight ways in which experiments studying the effect of iconicity on communication can correct for these confounds.

- Fay, N., Arbib, M., & Garrod, S. (2013). How to bootstrap a human communication system. *Cognitive Science*, *37*(7), 1356–67.
- Fay, N., Garrod, S., Roberts, L., & Swoboda, N. (2010). The interactive evolution of human communication systems. *Cognitive Science*, *34*, 351–386.
- Little, H., Eryılmaz, K., & Boer, B. de. (2017). Conventionalisation and discrimination as competing pressures on continuous speech-like signals. *Interaction studies*, 18(3), 355–378.
- Micklos, A. (2017). Iconic strategies in silent gesture: Perceiving the distinction between nouns and verbs. In *11th international symposium on iconicity in language and literature* (pp. 26–27).
- Occhino, C., Anible, B., Wilkinson, E., & Morford, J. P. (2017). Iconicity is in the eye of the beholder. *Gesture*, *16*(1), 100–126.
- Ortega, G., Schiefner, A., & Ozyurek, A. (2017). Speakers gestures predict the meaning and perception of iconicity in signs. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. Davelaar (Eds.), *the 39th annual conference of the cognitive science society (cogsci 2017)* (p. 889-894).
- Perlman, M., Dale, R., & Lupyan, G. (2015). Iconicity can ground the creation of vocal symbols. *Royal Society open science*, 2(8), 150152.
- Perry, L. K., Perlman, M., & Lupyan, G. (2015). Iconicity in english and spanish and its relation to lexical category and age of acquisition. *PloS one*, 10(9), e0137147.
- Sulik, J. (2018). Cognitive mechanisms for inferring the meaning of novel signals during symbolisation. *PLoS ONE*.
- Tamariz, M., Roberts, S. G., Martínez, J. I., & Santiago, J. (2017). The interactive origin of iconicity. *Cognitive Science*.
- Verhoef, T., Kirby, S., & Boer, B. (2016). Iconicity and the emergence of combinatorial structure in language. *Cognitive science*, 40(8), 1969–1994.

PHYSICAL FEATURES OF SOCIAL NETWORKS AND LANGUAGE TYPOLOGY

Matthew Lou-Magnuson*1 and Luca Onnis1

*Corresponding Author: matt0020@e.ntu.edu.sg ¹Nanyang Technological University, Singapore

Linguists have long noted that rich morphological patterns tend to appear in languages spoken by small groups more than larger ones (Evans & Levinson, 2009), and some have suggested that smaller social groups are simply better at supporting the kinds of innovation that lead to these developments (Trudgill, 2011; Nettle, 2012). In addition, languages seem to favor syntactic means over morphological ones as their communities of speakers grow in size. Indeed, empirical evidence suggests the typological patterning that languages display may be connected to aspects of the social network of the speakers. A recent survey of the World Atlas of Language Structures (WALS) (Dryer & Haspelmath, 2013) found that after controlling for phylogenetic and areal influence, a novel measure of population spread was highly correlated with the number of grammatical features marked by morphological constitutionality (Lupyan & Dale, 2010). Specifically, languages with smaller and more isolated speaker populations tend to make much greater use of morphology than those with larger and more wide-spread populations.

In order to investigate mechanisms that underly this correlation between social structure and language structure, we implemented a simple meaning and signal system (Spike, Stadler, Kirby, & Smith, 2017), integrating key features identified by linguists working on **grammaticalization**, the process responsible for structural change in language (Hopper & Traugott, 2003; Heine & Kuteva, 2002). In particular we tracked the number of times a signal was **reanalyzed** during intergenerational transfer, a necessary condition for the development of deeply nested morphological forms (Bybee, Perkins, & Pagliuca, 1994; Fortescue, 2016).

The model was run in two distinct experiments, both simulating repeated intergenerational transfer as relevant for the cultural evolution of language (Smith, Brighton, & Kirby, 2003). The first identified **transitivity** (aka the global clustering coefficient) as a physical property of social networks that closely approximates features put forth by sociolinguists as fundamental for morphological development (Trudgill, 2011; Nettle, 2012). We found that transitivity had a non-linear, positive correlation with a network's ability to support reanalysis, and further, that



Figure 1. Mean Level of Reanalysis vs Network Topology

there was a threshold below which a network was unable to support continued development.

The second experiment, building on the finding above, constructed four networks designed to closely mimic real world social structures. The results are summarized in Figure 1. The two networks, Random and Barabasi-Albert were present as controls, while Hierarchical was designed to mimic large and developed societies (Ravasz & Barabási, 2003), and Complete to mimic smaller, traditional ones. While both Complete and Hierarchical have transitivity levels above the threshold mentioned above, Hierarchical only supports reanalysis levels identical to Random, which has a transitivity far below the threshold. On examining the flow of signals in the network, we discovered that the hub agents that interconnected clusters served as developmental bottlenecks on the language. This hub pattern is characteristic of the **scale-free** property of complex networks, and a feature observed to develop in all modern human social networks, emerging once they become large enough, or possess internal hierarchies (Newman, 2010; Barabási, 2014).

Our findings provide an initial holistic and mechanistic account of the long standing correlation between social structure and language complexity. We propose transitivity as physical property of social networks responsible for supporting repeated gains in morphological complexity, and scale-freeness as a physical property responsible for suppressing and reversing such developments. Finally, beyond purely theoretical interests, the diversity of natural languages is undergoing unprecedented rates of attrition. If transitivity is a necessary component for the support of complex languages, then this metric can be applied to the social network of speakers attempting to revitalize minority languages.

References

Barabási, A.-L. (2014). Network science book. Network Science.

- Bybee, J., Perkins, R., & Pagliuca, W. (1994). *The evolution of grammar: Tense, aspect, and modality in the languages of the world.* University of Chicago Press.
- Dryer, M. S., & Haspelmath, M. (Eds.). (2013). *Wals online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and brain sciences*, 32(05), 429–448.
- Fortescue, M. D. (2016). Polysynthesis: a diachronic and typological perspective. *Oxford Research Encyclopedia (linguistics)*.
- Heine, B., & Kuteva, T. (2002). *World lexicon of grammaticalization*. Cambridge University Press.
- Hopper, P. J., & Traugott, E. C. (2003). *Grammaticalization*. Cambridge University Press.
- Lupyan, G., & Dale, R. (2010). Language structure is partly determined by social structure. *PloS one*, *5*(1), e8559.
- Nettle, D. (2012). Social scale and structural complexity in human languages. *Phil. Trans. R. Soc. B*, 367(1597), 1829–1836.
- Newman, M. (2010). Networks: an introduction. Oxford university press.
- Ravasz, E., & Barabási, A.-L. (2003). Hierarchical organization in complex networks. *Physical Review E*, 67(2), 026112.
- Smith, K., Brighton, H., & Kirby, S. (2003). Complex systems in language evolution: the cultural emergence of compositional structure. Advances in Complex Systems, 6(04), 537–558.
- Spike, M., Stadler, K., Kirby, S., & Smith, K. (2017). Minimal requirements for the emergence of learned signaling. *Cognitive science*, *41*(3), 623–658.
- Trudgill, P. (2011). Sociolinguistic typology: Social determinants of linguistic complexity. Oxford University Press.

CORE KNOWLEDGE OR LANGUAGE-AUGMENTED COGNITION? THE CASE OF GEOMETRIC REASONING

GARY LUPYAN^{*1}, ASHLEY WENDORF¹, LUIS MIGUEL BERSCIA,² JING PAUL³

*Corresponding Author: lupyan@wisc.edu
 ¹⁻²Department of Psychology, University of Wisconsin-Madison, USA
 ³Radboud University; MPI for Psycholinguistics, Nijmegen, Netherlands
 ⁴Other Department, Agnes Scott College, Decateur, GA USA

Understanding why language evolved in the hominid lineage requires understanding both the communicative as well as the cognitive functions of language. To what extent does language enable certain cognitive operations? What aspects of what we take to be 'normal' human cognition are enabled or facilitated by language? Some have argued that the answer is essentially 'none', and that while language is important for sharing and thereby accumulating cultural knowledge, our core cognitive processes are language-independent (e.g., Bloom & Keil, 2001; Gleitman & Fisher, 2005; Hespos & Spelke, 2004). Indeed, a dominant position within cognitive science posits that it is *because* humans are endowed with certain cognitive powers such as systematicity and symbolic reasoning that language learning and use is possible in the first place (e.g., Fodor, 2001; Laurence & Margolis, 1999).

On an alternate view, human cognition is transformed by the learning and use of a natural language (Bowerman & Levinson, 2001; Clark, 1998; Lupyan, 2012). In particular, language has been argued to be centrally involved in "nonverbal" tasks that require grouping together diverse exemplars based on specific criteria (Lupyan, 2009)—an essential component of symbolic thought. To the extent that this alternate view is correct, theories of language evolution need to expand their focus beyond the use of language for communication.

In this work we provide a strong test for the involvement of language in reasoning by examining geometric reasoning, a domain claimed to tap into a universal and language-independent human competence. Key evidence in support of universality and language-independence of geometric reasoning comes from a study by Dehaene et al. (2006) showing a strong correlation in performance on an odd-one-out geometric reasoning task between educated Americans and the Mundurukú, an Amazonian indigenous people without formal education and who do not possess vocabulary for describing the geometric relations in question.

We tested for causal involvement of language in geometric reasoning in six ways: (1) We examined whether overtly naming geometric relations improved

reasoning performance in English-speaking adults. (2) We examined whether interfering with language (having participants repeat "a b c" while doing the task for some trials) impaired performance. (3) We tested whether the ease of naming the relations in English selectively predicts English-speakers performance. (4) We tested whether ease of naming predicted the detrimental impact of verbal interference. (5) We tested congenitally deaf children residing in a Chinese special school for the deaf who were deprived of normal language input for most of their childhood, and compared their performance to children with normal language input and Chinese adults. (6) We collected data on the task from the Shawi, an indigenous group of horticulturalist traders from Northwestern Amazonia who speak a Kawapanan language. The Shawi we tested vary in formal education and knowledge of Spanish.

Our results replicate Dehaene et al's (2006) finding of substantial correlations in performance (r>.6) even among these very disparate populations. However, these correlations appear to reflect shared visual processing mechanisms rather than shared geometric reasoning abilities. We found strong evidence of causal involvement of language as revealed by: (1) overt naming improving performance (t=3.7), (2) verbal interference impairing performance (t=2.76), (3) significant correlations between performance and nameability (r>.5). (4) The effects of linguistic manipulations dependent on nameability: overt naming raised performance hard-to-name items (r=-.37); verbal interference selectively impaired performance on the normally easy-to-name items (r=.35). (5) Children with impaired language input performed substantially worse (M=.50) than children with normal language input (M=.75; t=4.1). The performance of the former was predicted by proficiency with Chinese sign language. (6) The Shawi performed poorly (M=.41) though, like the Mundurukú, considerably above chance (chance=.17). The Shawi's performance was strongly modulated by knowledge of Spanish. Importantly, neither the Mundurukú's nor the Shawi's responses were predicted by either English or Chinese nameability, suggesting that geometric relations that are easy to name are not universally accessible, but *become* easy when compact verbal descriptions are available.

Language may facilitate geometric reasoning by helping to construct a more categorical hypothesis space. Faced with an array of objects, the ability to name the objects and their relations (e.g., square, parallel, right-angle) provides an effective means of abstracting away perceptual factors that otherwise dominate the categorization response. The implications of this work reach beyond geometry showing how aspects of cognition frequently claimed to nonlinguistic in fact depend on linguistic enculturation and active language use.

References

- Bloom, P., & Keil, F. C. (2001). Thinking through language. *Mind & Language*, *16*(4), 351–367.
- Bowerman, M., & Levinson, S. C. (2001). Introduction. In M. Bowerman & S. C. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 1–16). Cambridge, UK: Cambridge University Press.
- Clark, A. (1998). Magic words: How language augments human computation. In P. Carruthers & J. Boucher (Eds.), *Language and Thought: Interdisciplinary themes* (pp. 162–183). New York, NY: Cambridge University Press.
- Dehaene, S., Izard, V., Pica, P., & Spelke, E. S. (2006). Core Knowledge of Geometry in an Amazonian Indigene Group. *Science*, 311(5759), 381–384. https://doi.org/10.1126/science.1121739
- Fodor, J. A. (2001). Language, Thought and Compositionality. *Mind & Language*, *16*(1), 1–15. https://doi.org/10.1111/1468-0017.00153
- Gleitman, L., & Fisher, C. (2005). Universal aspects of word learning. In J. McGilvray (Ed.), *The Cambridge Companion to Chomsky* (pp. 123–144). New York, NY: Cambridge University Press. Retrieved from https://books.google.com/books?hl=en&lr=&id=I6CZ6wpNKeEC&oi=fnd& pg=PA123&ots=ujaHwtyK8O&sig=N-te0BeTbB0IyV7aOmuwBAqwtCk
- Hespos, S. J., & Spelke, E. S. (2004). Conceptual precursors to language. *Nature*, 430(6998), 453–456.
- Laurence, S., & Margolis, E. (1999). Concepts and Cognitive Science. In E. Margolis & S. Laurence (Eds.), *Concepts: Core Readings* (pp. 3–81). Cambridge, Mass: A Bradford Book.
- Lupyan, G. (2009). Extracommunicative Functions of Language: Verbal Interference Causes Selective Categorization Impairments. *Psychonomic Bulletin & Review*, 16(4), 711–718. https://doi.org/10.3758/PBR.16.4.711
- Lupyan, G. (2012). What do words do? Towards a theory of languageaugmented thought. In B. H. Ross (Ed.), *The Psychology of Learning and Motivation* (Vol. 57, pp. 255–297). Waltham, MA: Academic Press. Retrieved from

http://www.sciencedirect.com/science/article/pii/B9780123942937000078

GESTURE SIGN AND BEYOND NEGATION ACROSS THREE GENERATIONS OF SIGNERS

Hannah Lutzenberger^{*1,2}

*Corresponding Author: h.lutzenberger@let.ru.nl ¹Center of Language Studies, Radboud University, Nijmegen, The Netherlands ² International Max Planck Research School, Nijmegen, The Netherlands

Unlike spoken languages for which written records date back several millennia, village sign languages emerge from scratch over the timespan of only a few decades even in the 20th century. These languages newly arise in remote areas where there is an exceptionally high incidence of hereditary deafness (Zeshan & de Vos, 2012). Kata Kolok (KK) is a young sign language used by hearing and deaf members of a single Balinese village community (Marsaja, 2008). The present study investigates negation in KK.

Negation is a core feature of all human languages, both spoken and signed (Dahl, 2010; Miestamo, 2005). Across a large range of unrelated signed languages, negation is expressed by both manual and non-manual elements: a lexical sign with a repeated side-to-side movement and a headshake (Zeshan, 2004). Often shared with the wider hearing community, the manual and the non-manual marker seem to have derived from co-speech gestures (Wilcox, 2009). In sign languages, these gestures have grammaticalized into linguistic negation markers that vary alongside two parameters: 1) only one of the two markers (manual or non-manual) is obligatory, and 2) spreading of the headshake i.e. the headshake is co-articulated only with the manual negator or extends to neighboring signs.

I study negation in spontaneous conversations from six KK signers of the generations III through V (KK Corpus; de Vos, 2016). Transcription of 1.73 hours of data reveals two main findings: (i) Across all three generations, KK signers rely on formally identical negation markers that originate in conversational and cultural gestures (Marsaja, 2008): a lexicalized handwave (manual negator), a negative headshake, and tongue protrusion. Whilst the latter occurs in specialized contexts of negative evaluation and negative existence, the

data did not unambiguously identify a single obligatory negation marker. The manual negator **and** the headshake are highly frequent in all signers, both combined and independent. This makes KK negation typologically unusual. (ii) Inter-generational differences in the degree of headshake spreading indicate language change across three generations: a linear mixed-effects model reveals that the headshake extends to neighboring signs in the youngest generation of signers, but signers from older generations almost always co-produce the headshake only with the manual negator (Fig. 1).



Figure 1. Significant intergenerational difference in headshake spreading.

Headshake spreading is thought to lead to more efficient language use since using the non-manual headshake as mandatory negation marker frees the hands for other signs (Pfau, 2015). Sign languages that opt for this multi-channelled pattern have a higher compressibility than sign languages that require the manual negator to negate. This non-manual system is attested in most sign languages studied that are considerably older than KK (Zeshan, 2004). Headshake spreading in KK, however, does not occur alongside a decrease in the manual negator. Thus, the hand is actually not available for other signs. It is possible that KK exploits two obligatory negation markers - a pattern that has not yet been described. An alternative explanation might be sought in the time depth of the language: the youngest generation of signers might represent a transitional state of the developing negation system. In other words, the patterns of negation found in KK may not yet have been attested in another language because of its exceptional emergence scenario.

The results from this study show that even within a relatively short timeframe, emergent sign languages like KK may evolve to have unique and previously unattested linguistic features and thus provide a window to the study of the beginnings of modern human languages (de Vos & Pfau, 2015).

- Dahl, Ö. (2010). Typology of negation. In L. R. Horn (Ed.), *The expression of negation*. (pp. 9-38). Berlin: Mouton de Gruyter.
- de Vos, C. (2016). Sampling shared sign languages. Sign Language Studies 16(2), 204-226.
- De Vos, C., & Pfau, R. (2015). Sign Language Typology: The contribution of rural sign languages. Annual Review of Linguistics, 1, 265-288.
- Marsaja, G. I. (2008). Desa kolok: A deaf village and its sign language in Bali, Indonesia. Nijmegen: Ishara Press.
- Miestamo, M. (2005). *Standard negation. The negation of declarative verbal main clauses in a typological perspective.* Berlin: Mouton de Gruyter.
- Pfau, R. (2015). The grammaticalization of headshakes: From head movement to negative head. In A. D. M. Smith, G. Trousdale & R. Waltereit (Eds.), *New directions in grammaticalization research* (pp. 9-50). Amsterdam: John Benjamins.
- Wilcox, S. (2009). Symbol and symptom: Routes from gesture to signed language. *Annual Review of Cognitive Linguistics* 7, 89-110.
- Zeshan, U. (2004). Hand, head and face Negative constructions in sign languages. *Linguistic Typology* 8 (1), 1-58.
- Zeshan, U., & de Vos, C. (Eds.) (2012). Sign languages in village communities: Anthropological and linguistic insights. Berlin: De Gruyter Mouton & Nijmegen: Ishara Press.

SYMBOL AND ITS EVOLUTION

ERKKI LUUK

erkkil@gmail.com Institute of Computer Science, University of Tartu, Estonia

The paper investigates the problem of tracing the emergence of symbolism in the *Homo* lineage. We define the symbol via the notions of arbitrary and spatiotemporally displaced reference, and analyze the earliest manifestations of different forms of symbolism: color, figurative, abstract (in the form of codes, other signs, and ritual burials), and (proto)language. Besides symbols themselves, diverse physical and behavioral traits that might constitute a circumstantial evidence for symbolism are scrutinized. Drawing on archaeological and fossil evidence, a plausible time for the emergence of symbol in protolanguage is estimated to be 850 ... 2200 kya.

1. Symbol

1.1. Introduction

Symbolism is almost universally viewed as a distinctively (frequently even exclusively) human trait, and one that is fundamental to human language. Owing to the multi-faceted nature of symbol (more of which below), the scarcity of direct evidence of symbolism from Middle and Lower Paleolithic (c. 50 ... 300 and 300 ... 3300 kya, respectively), and the uncertainty of circumstantial evidence from rare fossil and (even rarer) archaeological finds that might imply symbolism, tracing its emergence in the human lineage is a complex task.

1.2. Definition

As 'symbol' is a pretheoretic term, it has to be defined from the outset. Two features that are usually, if not implicitly, held to be characteristic of 'symbol' are the arbitrary nature of reference (a non-necessary link between a sign and its meaning) and a used potential for spatiotemporally displaced reference¹. Arbi-

Abbreviations: kya - thousand years ago.

¹The above definition of symbol opens the possibility that, in natural communication, symbolism is uniquely human. For example, while bottlenose dolphins' signature whistles are arbitrary, the whistles' natural potential for spatiotemporally displaced reference is not evident, as each dolphin uses its own distinctive signature whistle (Janik, Sayigh, & Wells, 2006). Although dolphins frequently copy each other signature whistles in the wild, it is not clear whether this qualifies as a reference to third person individuals (which would indicate a used potential for spatiotemporally displaced reference).

trariness is closely related to the number of signs (the more signs the greater the pressure for their arbitrariness). In the form of grammaticalization, the prevailing drift of natural languages towards arbitrariness is evident both on historical and evolutionary scales (Heine & Kuteva, 2007).

Differently from symbols, 'index' and 'icon' are defined by a necessary and natural link between a sign and its meaning. In icons, the link is based on similarity. In indices, it is based on any other necessary relation (e.g. cause-effect or whole-part relation). In practice, various degrees of iconicity as possible, and symbols (e.g. the Christian cross, or onomatopoetic words like *knock*, *bump*, *crash*) may exhibit iconicity as well (Luuk, 2013). Equipped with these definitions, one notices that the word 'symbol' gets colloquially used in very different senses. For example, status symbols (e.g. words). By the above definitions, only the latter are symbols. The former are indices of wealth and, more generally, success. Although the difference might seem minor, it has fundamental implications on the archaeological evidence for symbolism. For example, one cannot infer symbolism (and by extension, language) from personal ornaments, as the most parsimonious interpretation of personal ornaments is that they are status symbols (Sterelny, 2008).

2. Ornamental symbolism

As personal ornaments are costly, their are indices of success. This observation is archaeologically supported by the fact that most preserved² objects used in Pleistocene pendants (e.g. predator teeth and suitable seashells) are hard to come by. As killing a predator is evolutionarily extremely costly, it is a bona fide indicator of success. Nevertheless, it is exceedingly common to interpret personal ornaments (and to a lesser extent pigment use) as manifestations of symbolism (e.g. Bednarik, 2008b; Chase, 1994; d'Errico et al., 2003; d'Errico & Backwell, 2016; Langley, O'Connor, & Piotto, 2016). Observe also that personal ornaments do not imply displaced reference, as they bestow status only to their wearers. Thus, personal ornaments *per se* are not indicators of symbolism (and, by extension, language). Of course, an indexical function does not preclude a symbolic one (and objects can be clearly symbolic as well — we will see some examples below) but non-symbolism should be the null hypothesis for objects with an indexical function.

However, even non-symbolic objects may belong to a behavioral pattern, some other parts of which are indicative of symbolism. For example, a procurement and processing of beads, pendants or colorants for using them in body or garment ornamentation puts up a much stronger case for symbolism than the (processed)

²Some well-preserved Holocene burials have contained personal ornaments of perishable materials, e.g. seeds beads (Tryon & Faith, 2013).

objects themselves. The evidence gets stronger if the nearest source for the objects lies at least several kilometres from the place they were processed and/or applied. Of course, the evidence is circumstantial, but it constitutes a much stronger case for symbolism than when possibly non-symbolic objects are just claimed to be symbols.

3. Color symbolism

The earliest evidence of colorant processing - ochre on grindstones - is from the Kapthurin formation in Kenya (284 ... 500 kya, Tryon & Faith, 2013). In general, it is very difficult to rule out all potential utilitarian uses for Middle and Lower Paleolithic colorant finds (medicine, hide preservative, protection from sun and insects, camouflage, startling of prey and conspecifics, the preparation of mastics for hafting, etc. — Barham, 2002; Sterelny, 2008; Wadley, 2001; Wadley, Williamson, & Lombard, 2004) but even where this could be done, there remains the possibility that pigment was used because definitive colors were preferred for aesthetic or cognitive (salience) reasons. Even non-human species differentiate between aesthetic and non-aesthetic stimuli and utilize definitive colors as behavioral cues (Watanabe, 2010) and so do children in their first year (Baldwin, 2006). While coloring is probably uniquely human, there is nothing inherently symbolic about it. For color symbolism to be present, a non-natural, non-random and nonavailability-specific link between color and object (or color and figure) has to be evidenced. For example, a brown, black or white foot figure on a cave wall is probably color iconism (reference by similarity), one purple foot figure is probably a chance but ten purple foot figures suggest color symbolism (except if purple was one of the few pigments available to the artist and the others were equally non-iconic, e.g. crimson and green). Thus, extremely specific configurations of archeological and geological evidence are required to attest color symbolism.

4. Figurative symbolism

Figurative symbolism is generally easier to establish. Importantly, as figurative paintings and sculptures are at least partly iconic, representational art per se does not entail symbolism (Luuk, 2013). Thus, the vast majority of cave paintings and early sculptures (including the Berekhat Ram figurine from 250 kya — d'Errico et al., 2003) have to be excluded from possible indicators of symbolism. However, some early examples of figurative symbolism remain. For example, the two half-lion/half-man figurines from 31 kya (Conard, 2003) are symbolic, as they exhibit both spatiotemporally displaced and arbitrary reference (spatiotemporally displaced, because a half-lion/half-man has hardly any potential to refer to any-thing here and now, and arbitrary because the figurines are not representational as wholes). By extension, the figurines are reasonable proxies for language (or at least protolanguage). A similar case could be made about the Hohle Fels fe-

male figurine (Conard, 2009), which, however, would be better analyzed in the following section.

5. Abstract symbolism

In abstract symbolism, an important distinction is between codes and elementary abstract signs. Only the former can be compositional. The codes may represent calendars, bookkeeping tables, written texts, etc. Another difference is that codes are clearly symbolic (with elementary abstract signs this may not be so clear).

5.1. Codes

Very few examples of codes from (Upper) Paleolithic have been found³. One of the first examples of complex code appears on the La Marche antler from 16 kya (d'Errico, 1995). Simple codes are probably in evidence since at least the Ishango bone from c. 20 kya (Bogoshi, Naidoo, & Webb, 1987; Brooks & Smith, 1987). No preserved examples of codes from Lower and Middle Paleolithic are known.

5.2. Signs

The majority of putative abstract "signs" from Middle and Lower Paleolithic fall into a category of 'abstract art' for which doodling is a more parsimonious explanation than symbolism (Bednarik, 1995a; Halverson, 1995). For example, Bednarik's (1995a) paper on concept-mediated marking in Lower Paleolithic makes no allegations as to the symbolicity or even intentionality of the markings. The markings that he analyzed were abstract, and some of them were very similar to those Middle Paleolithic markings found on bone and ochre pieces from Blombos Cave that are claimed to be "irrefutable evidence of symbolic behavior" (d'Errico et al., 2003, p. 4). The claim relies, of course, on a pretheoretic notion of symbol. However, some early examples of abstract symbolism remain.

Markings on the above-mentioned ivory female figurine from Hohle Fels, Germany, dated to the Aurignacian 35 kya, have been interpreted as ideograms representing the Upper Paleolithic Double Goddess (Harrod, 2011). While the interpretation may seem dubious, the distorted and geometrically decorated figurine is very likely symbolic.

An impressive collection of abstract signs have been discovered in Chauvet cave in France (Petzinger & Nowell, 2014), the earliest of which are dated to 35 ... 38.5 kya (early Upper Paleolithic). The number (17) of distinctive non-figurative signs at Chauvet is striking. However, by being scattered spatially (and likely also temporally), they do not form a code any more than other abstract art styles do. Nevertheless, the signs' abstract (albeit half-iconic — cruciform, reniform,

³None of them is definitely deciphred.

negative and positive hand, etc.) nature more or less qualifies them as symbols, even though one has no idea what they stood for.

Crossing over to Middle Paleolithic, a tradition of engraving ostrich egg shells, dated to 60 kya, has been discovered in South Africa (Texier et al., 2010). The fixed, abstract and persistent nature of the engravings qualifies them as symbols of (at least) group identity. No doubt to the late engravers the hatched band motive they reproduced for millenia seemed rather pregiven than 'arbitrary' (in another but a related meaning of the word). In addition, for those who started the tradition, the engravings may have been iconic (e.g. of a spun rope). However, the tradition's length ensures that they became abstract, symbolic and arbitrary (in the definitive meaning).

In general, cupules are better candidates for symbols than other forms of 'abstract art' because their manufacture is labor intensive, which rules out doodling as well as coincidental configurations of cut marks (Bednarik, 1995b, 2008a, 2008b). At the same time, it is plausible that some cupules (at least those on horizontal surfaces) were used as containers or were unintentional byproducts of other functional activities (e.g. grinding). Although a functional role does not preclude a symbolic use, it makes cupules' status as an evidence of symbolism ambiguous. Nevertheless, cupules are seemingly a later addition to hominins' behavioral repertoire than personal ornaments (at least 150 and 300 kya, respectively — Bednarik, 2008a, 2008b).

5.3. Burials

Although emotional attachment is a more parsimonious explanation for burials than symbolism (Sterelny, 2008), grave goods, structures and their configurations can point to symbolism as well. A Neanderthal burial site, La Ferrassie, dated to 65-70 kya, suggests a possible symbolic activity (Zilhão et al., 2010; Bednarik, 1995a, 2008a). The find that was perhaps the most implicative of symbolism in La Ferrassie was a limestone slab with 18 cupules (16 of which were arranged in pairs), covering a child's corpse (Bednarik, 2008a, citing Peyroni, 1934). The earliest modern human burial, dated to 74 kya, has been recently excavated in Border Cave in South Africa. The burial was associated with personal ornaments (perforated Conus shells — d'Errico & Backwell, 2016).

5.4. Language

By far the oldest evidence for symbolism we have is circumstantial and, ironically, it is an evidence for language, the most elusive form of symbolism archaeologically. But first, some qualifications. We are interested in the time the language evolved. By "language" we mean (proto)language, i.e. a language which could be a protolanguage. A rough definition of protolanguage would be "a human language with a nonmodern syntax". A precise one could be "a human language without cases and adpositions". The function of cases and adpositions is to mark arguments of the head of the sentence. The sentence is a precondition of any syntax, modern or nonmodern. In modern languages, the head of the sentence is usually a finite verb or flexible⁴. As no modern languages without cases or adpositions are attested⁵, they are a suitable dividing line. By the "time language evolved" we mean the time it emerged. Johanssen (2005) constrained this time to between 500 and 5000 kya, which is the necessary period the language had to emerge. We are interested in refining it, i.e. in a both necessary and plausible (with more than 0.5 probability) timeline of the evolution of language.

5.4.1. Evidence from colonization

It is well known that *Homo erectus*⁶ crossed at least 19 km of open sea to colonize Flores c. 850 kya or earlier (Gibbons, 1998a; Morwood, O'Sullivan, Aziz, & Raza, 1998). It has been (plausibly) argued that watercraft manufacture and navigation entail a level of communication that would be unattainable without a (proto)language. Although rafting on flotsam is a possibility (Gibbons, 1998a), it is not a likely one, given *H. erectus*'s ability to manufacture mode 2 tools and successful colonization of much of the Old World (from Africa and Western Europe to Java, China and, possibly, Central Siberia — Ascenzi, Benvenuti, & Segre, 1997; Asfaw et al., 2002; Bar-Yosef & Belfer-Cohen, 2001; Gibbons, 1998b; Larick et al., 2001; Waters, Forman, & Pierson, 1997).

For taphonomic reasons, a direct evidence of symbolic activity in so distant times would be a find of centuries. Likewise, the colonization of the southern arc by modern humans much later (likely 65 ... 90 kya — Clarkson et al., 2017) provides clear evidence of symbolic conceptualization but very few symbols before the colonizers arrived in Australia (Balme, Davidson, McDonald, Stern, & Veth, 2009). Even more surprisingly, there is a higher proportion of sites with symbols early in Australia's colonization rather than later (Balme et al., 2009).

5.4.2. Anatomical evidence

As compared to other extinct hominids, medium and late *H. erectus* (less than 1600 kya) had an increase in thoracic innervation, similar to that of modern humans and *H. Neanderthalis* (MacLarnon & Hewitt, 1999; Meyer, Lordkipanidze, & Vekua, 2006). The authors associate it with an enhanced breathing control, which most likely evolved to facilitate speech (MacLarnon & Hewitt, 1999). Fitch (2000) notes that it is difficult to know whether increased respiratory control directly involved speech, or evolved for other reasons (e.g. prolonged running or

⁴There is no general agreement as to whether 'verb' is in fact universal in modern languages (or how to define it — Himmelmann, 2007; Luuk, 2010).

⁵Pidgins, which are not full blown languages, are a possible exception.

⁶For simplicity, I assume the null hypothesis (proposed by Lordkipanidze et al., 2013) that *H. habilis, H. ergaster* and *H. rudolfensis* belong to a single evolving lineage of *H. erectus*.

swimming), and provided only a necessary preadaptation to speech. There is another aspect to this argument. By "speech" is meant modern speech, but no-one assumes that *H. erectus* was capable of that. By definition, a more modest range of vocalization that could support protolanguage would have been sufficient for its production.

An anatomical difference that allows humans to produce a wider range of formant patterns than other mammals is a lowered larynx (Lieberman, Klatt, & Wilson, 1969). Lieberman (1987) suggests that the lowering of larynx started with *H. erectus*⁷. Since the decent of larynx has an evolutionary cost (vulnerability to choking — Fitch, 2000), there must have been a selective pressure for the descent. In addition, to bestow a phonetic advantage, larynx must be significantly lower than its position in extant nonhuman apes (Fitch, 2000). At least two adaptive scenarios explain the presumably gradual descent of the larynx in *Homo*: facilitating mouth breathing during extreme physical challenge (Lieberman, 1987) and displays of exaggerated power⁸ to intimidate enemies and competitors and attract potential mates (Ohala, 1984). Both scenarios might result in the descent of larynx as a preadaptation to speech.

5.4.3. Conclusion

Three independent and converging pieces of evidence (1, 2, 3):

- *H. erectus*' colonization of Flores and much of the other Old World (Gibbons, 1998a; Morwood et al., 1998; Ascenzi et al., 1997; Asfaw et al., 2002; Bar-Yosef & Belfer-Cohen, 2001; Gibbons, 1998b; Larick et al., 2001; Waters et al., 1997),
- 2. *H. erectus*' enhanced respiratory control (MacLarnon & Hewitt, 1999; Meyer et al., 2006),
- 3. A lowering of larynx in H. erectus (Lieberman, 1987),
- 4. An enhanced capacity for vocalizations in H. erectus (2, 3),
- 5. An enhanced capacity for vocal communication in H. erectus (4),

point to some kind of language in *H. erectus* prior to its colonization of Flores. Of course, the evidence is circumstantial but this is always the case in language evolution — the earliest direct evidence for language (an inscription on a wooden tablet — Whitley, 2003) is from less than 10 kya.

 $^{^{7}}$ Larynx does not fossilize, so its position can only be conjectured from fossiles (and reliable clues are lacking — Fitch, 2000).

⁸A lowered larynx allows to imitate vocalizations of larger animals that lack this feature (Fitch, 2000).

As *Homo habilis*, the first known *Homo* species (with the earliest occurrence c. 2300 kya — Spoor et al., 2007), was not scrutinized by MacLarnon and Hewitt (1999), it is possible that this species, too, was capable of enhanced vocalizations. According to Tobias (1998), evidence from endocranial casts suggests that *H. habilis* was a "speaker"⁹. When combined, the archaeological and anatomical evidence indicate that *H. erectus* (possibly even starting from *H. habilis*) used some kind of (proto)language, the emergence of which could be tentatively bracketed between 850 and 2200 kya. At present, this constitutes the earliest plausible evidence of symbolism. As mentioned above, the time when language evolved was previously constrained to 500 ... 5000 kya (Johansson, 2006).

6. Conclusion

In this paper we have (1) defined the symbol, (2) analyzed some possible proxies of symbolism, and (3) reconstructed a plausible timeline for the emergence of symbol (viz., in protolanguage) by drawing on relevant fossil and archaeological evidence. We have bracketed the plausible time of the emergence of symbol to 850 ... 2200 kya. Symbol is defined by its two distinctive features: (a) an arbitrary nature (a non-necessary link between a sign and its meaning), and (b) a used potential for spatiotemporally displaced reference. The traits we have analyzed include behavioral patterns (e.g. coloring and engraving traditions associated with procurement and processing of materials, watercraft manufacture and navigation, ritual burials associated with personal ornaments, etc.) and anatomical features like thoracic innervation and the position of larynx. To a different degree, all these features can be viewed as proxies for symbolism. Without a doubt, an engraving tradition makes a much better proxy for symbolism than an extinct species' (possibly) lowered larynx, but it is also a much more recent one (on the order of 100 and 1000 kya, respectively). For taphonomic reasons, fossil evidence dominates the >500 kya era over the archaeological, thus being the main window to the behavioral capacities of the hominids of the period.

Acknowledgements

This work has been supported by IUT20-56 and European Regional Development Fund through CEES.

⁹It has been also proposed that language originated in gestural modality (Corballis, 2010). This is supported by the observation that our closest extant relatives are proficient in gestural but poor at vocal imitation (Pika & Mitani, 2006). If the scenario is true, the emergence of language might have predated a selection for enhanced vocalizations. We have not considered this possibility in reconstructing the present timeline.

- Ascenzi, A., Benvenuti, A., & Segre, A. G. (1997). On the paleopathologic findings exhibited by the late *Homo erectus* of Ceprano, Italy. *Human Evolution*, 12(3), 189–196.
- Asfaw, B., Gilbert, W. H., Beyene, Y., Hart, W. K., Renne, P. R., WoldeGabriel, G., Vrba, E. S., & White, T. D. (2002). Remains of *Homo erectus* from Bouri, Middle Awash, Ethiopia. *Nature*, 416, 317–320.
- Baldwin, J. M. (2006). Handbook Of Psychology: Senses And Intellect (2nd [orig. 1890] ed.). Whitefish, MT: Kessinger Publishing.
- Balme, J., Davidson, I., McDonald, J., Stern, N., & Veth, P. (2009). Symbolic behaviour and the peopling of the southern arc route to Australia. *Quaternary International*, 202(1), 59 - 68. (Great Arc of Human Dispersal)
- Barham, L. S. (2002). Systematic pigment use in the Middle Pleistocene of South-Central Africa. *Current Anthropology*, 43(1), 181–190.
- Bar-Yosef, O., & Belfer-Cohen, A. (2001). From Africa to Eurasia early dispersals. *Quaternary International*, 75(1), 19–28.
- Bednarik, R. G. (1995a). Concept-mediated marking in the Lower Palaeolithic. *Current Anthropology*, *36*(4), 605–616.
- Bednarik, R. G. (1995b). Reply to comments on Bednarik's "Concept-mediated marking in the Lower Palaeolithic". *Current Anthropology*, *36*(4), 626–630.
- Bednarik, R. G. (2008a). Cupules. Rock Art Research, 25(1), 61-100.
- Bednarik, R. G. (2008b). The origins of symboling. Signs, 2, 82-113.
- Bogoshi, J., Naidoo, K., & Webb, J. (1987). The oldest mathematical artifact. *Mathematical Gazette*, 71, 458.
- Brooks, A. S., & Smith, C. C. (1987). Ishango revisited: new age determinations and cultural interpretations. *African Archaeological Review*, 5(1), 65–78.
- Chase, P. G. (1994). On symbols and the Palaeolithic. *Current Anthropology*, 35(5), 627–629.
- Clarkson, C., Jacobs, Z., Marwick, B., Fullagar, R., Wallis, L., Smith, M., Roberts, R. G., Hayes, E., Lowe, K., Carah, X., Florin, S. A., McNeil, J., Cox, D., Arnold, L. J., Hua, Q., Huntley, J., Brand, H. E. A., Manne, T., Fairbairn, A., Shulmeister, J., Lyle, L., Salinas, M., Page, M., Connell, K., Park, G., Norman, K., Murphy, T., & Pardoe, C. (2017). Human occupation of northern Australia by 65,000 years ago. *Nature*, 547(7663), 306–310.
- Conard, N. J. (2003). Palaeolithic ivory sculptures from southwestern Germany and the origins of figurative art. *Nature*, 426, 830–832.
- Conard, N. J. (2009). A female figurine from the basal Aurignacian of Hohle Fels Cave in southwestern Germany. *Nature*, *459*(7244), 248–252.
- Corballis, M. C. (2010). The gestural origins of language. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(1), 2–7.
- d'Errico, F. (1995). A new model and its implications for the origin of writing:

La Marche antler revisited. *Cambridge Archaeological Journal*, 5(2), 163–206.

- d'Errico, F., & Backwell, L. (2016). Earliest evidence of personal ornaments associated with burial: The *Conus* shells from Border Cave. *Journal of Human Evolution*, 93(Supplement C), 91 108.
- d'Errico, F., Henshilwood, C., Lawson, G., Vanhaeren, M., Tillier, A.-M., Soressi, M., Bresson, F., Maureille, B., Nowell, A., Lakarra, J., Backwell, L., & Julien, M. (2003). Archaeological evidence for the emergence of language, symbolism, and music an alternative multidisciplinary perspective. *Journal* of World Prehistory, 17(1), 1–70.
- Fitch, W. T. (2000). Vocal production in nonhuman mammals: implications for the evolution of speech. In *Evolution of Language Conference*. Paris.
- Gibbons, A. (1998a). Ancient island tools suggest *Homo erectus* was a seafarer. *Science*, 279(5357), 1635–1637.
- Gibbons, A. (1998b). In China, a handier Homo erectus. Science, 279(5357), 1636.
- Halverson, J. (1995). Comment on Bednarik's "Concept-mediated marking in the Lower Palaeolithic". *Current Anthropology*, *36*(4), 621.
- Harrod, J. B. (2011). The Hohle Fels female figurine: Not pornography but a representation of the Upper Paleolithic Double Goddess. *Journal of Archaeomythology*, 7, 204–218.
- Heine, B., & Kuteva, T. (2007). *The genesis of grammar: a reconstruction*. New York: Oxford University Press.
- Himmelmann, N. P. (2007). Lexical categories and voice in Tagalog. In P. Austin & S. Musgrave (Eds.), *Voice and Grammatical Functions in Austronesian Languages.* Stanford: CSLI.
- Janik, V. M., Sayigh, L. S., & Wells, R. S. (2006). Signature whistle shape conveys identity information to bottlenose dolphins. *Proceedings of the National Academy of Sciences of the United States of America*, 103(21), 8293–7. (Journal Article Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, Non-P.H.S.)
- Johansson, S. (2006). Constraining the time when language evolved. In A. Cangelosi, A. D. M. Smith, & K. Smith (Eds.), *The Evolution of Language: Proceedings of the 6th International Conference (EVOLANG6)* (pp. 152– 159). Singapore: World Scientific.
- Langley, M. C., O'Connor, S., & Piotto, E. (2016). 42,000-year-old worked and pigment-stained *Nautilus* shell from Jerimalai (Timor-Leste): Evidence for an early coastal adaptation in ISEA. *Journal of Human Evolution*, 97(Supplement C), 1 - 16.
- Larick, R., Ciochon, R. L., Zaim, Y., Sudijono, Suminto, Rizal, Y., Aziz, F., Reagan, M., & Heizler, M. (2001). Early Pleistocene 40ar/39ar ages for Bapang Formation hominins, Central Jawa, Indonesia. *Proceedings of the National*
Academy of Sciences of the United States of America, 98(9), 4866–4871.

- Lieberman, P. I. (1987). *The Biology and Evolution of Language* (2nd [orig. 1984] ed.). Harvard University Press.
- Lieberman, P. I., Klatt, D. H., & Wilson, W. A. (1969). Vocal tract limitations on the vowel repertoires of rhesus monkey and other nonhuman primates. *The Journal of the Acoustical Society of America*, 46(1A), 97-97.
- Lordkipanidze, D., León, M. S. Ponce de, Margvelashvili, A., Rak, Y., Rightmire, G. P., Vekua, A., & Zollikofer, C. P. E. (2013). A complete skull from Dmanisi, Georgia, and the evolutionary biology of early *Homo. Science*, 342(6156), 326–331.
- Luuk, E. (2010). Nouns, verbs and flexibles: implications for typologies of word classes. *Language Sciences*, 32(3), 349–365.
- Luuk, E. (2013). The structure and evolution of symbol. *New Ideas in Psychology*, *31*, 87–97.
- MacLarnon, A. M., & Hewitt, G. P. (1999). The evolution of human speech: The role of enhanced breathing control. *American Journal of Physical Anthropology*, 109, 341–363.
- Meyer, M. R., Lordkipanidze, D., & Vekua, A. (2006). Evidence for the anatomical capacity for spoken language in *Homo erectus*. *American Journal of Physical Anthropology Supplement*, 42, 130.
- Morwood, M. J., O'Sullivan, P. B., Aziz, F., & Raza, A. (1998). Fission-track ages of stone tools and fossils on the east Indonesian island of Flores. *Nature*, *392*, 173–176.
- Ohala, J. J. (1984). An ethological perspective on common cross-language utilization of F0 of voice. *Phonetica*, 41(1), 1-16.
- Petzinger, G. von, & Nowell, A. (2014). A place in time: Situating Chauvet within the long chronology of symbolic behavioral development. *Journal* of Human Evolution, 74, 37 54.
- Peyroni, D. (1934). La Ferrassie. Préhistoire, 3, 1-92.
- Pika, S., & Mitani, J. (2006). Referential gestural communication in wild chimpanzees (*Pan troglodytes*). *Current Biology*, *16*(6).
- Spoor, F., Leakey, M. G., Gathogo, P. N., Brown, F. H., Antn, S. C., McDougal, I., Kiarie, C., Manthi, F. K., & Leakey, L. N. (2007). Implications of new early *Homo* fossils from Ileret, east of Lake Turkana, Kenya. *Nature*, 448, 688–691.
- Sterelny, K. (2008). What is Behavioural Modernity? *Paper for the Nicod Lecture*, http://www.institutnicod.org/Sterelny-WhatisBehaviouralMordernity.pdf.
- Texier, P.-J., Porraz, G., Parkington, J., Rigaud, J.-P., Poggenpoel, C., Miller, C., Tribolo, C., Cartwright, C., Coudenneau, A., Klein, R., Steele, T., & Verna, C. (2010). A howiesons poort tradition of engraving ostrich eggshell containers dated to 60,000 years ago at diepkloof rock shelter, south africa. *Proceedings of the National Academy of Sciences*, 107(14), 6180-6185.

- Tobias, P. V. (1998). Evidence for the early beginnings of spoken language. *Cambridge Archaeological Journal*, 8(1), 72–78.
- Tryon, C. A., & Faith, J. T. (2013). Variability in the Middle Stone Age of Eastern Africa. *Current Anthropology*, *54*(S8), S234-S254.
- Wadley, L. (2001). A general view and a South African perspective from Rose Cottage Cave. *Cambridge Archaeological Journal*, *11*(2), 201–221.
- Wadley, L., Williamson, B., & Lombard, M. (2004). Ochre in hafting in Middle Stone Age southern Africa: a practical role. *Antiquity*, 78, 661–675.
- Watanabe, S. (2010). Pigeons can discriminate good and bad paintings by children. *Animal Cognition*, *13*(1), 75–85.
- Waters, M. R., Forman, S. L., & Pierson, J. M. (1997). Diring Yuriakh: A Lower Paleolithic site in Central Siberia. *Science*, 275, 1281–1284.
- Whitley, J. (2003). Archaeology in Greece 2003-2004. *Archaeological Reports*, 50, 1–92.
- Zilhão, J., Angelucci, D. E., Badal-García, E., d'Errico, F., Daniel, F., Dayet, L., Douka, K., Higham, T. F. G., Martínez-Sánchez, M. J., Montes-Bernárdez, R., Murcia-Mascarós, S., Pérez-Sirvent, C., Roldán-García, C., Vanhaeren, M., Villaverde, V., Wood, R., & Zapata, J. (2010). Symbolic use of marine shells and mineral pigments by Iberian Neandertals. *Proceedings of the National Academy of Sciences*, 107(3), 1023-1028.

GREETING AND MEETING: VOCAL INTERACTIONS AND USE OF SIGNATURE WHISTLES IN BOTTLENOSE DOLPHINS (TURSIOPS TRUNCATUS) DURING AN INTRODUCTION

Heidi Lyn and Megan Broadway

*Corresponding Author: Heidi.lyn@usm.edu ¹Department of Psychology, University of Southern Mississippi, Long Beach, MS 39560

Bottlenose dolphins are the focus of a great deal of mythology that casts all dolphins and whales as on par or exceeding humans in their intelligence and communicative abilities (Gregg, 2013). Experiments with artificial language systems have shown that dolphins can attach meaning to symbols (Herman, 1987) and can follow arbitrary ordering rules (Herman, Kuczaj, & Holder, 1993), much like human syntax, but have failed to find elements that suggest the complexity of human language (Kako, 1999). However, dolphins have been found to utilize a rare call type that has only been found in humans, some parrots, and dolphins (Balsby & Bradbury, 2009; Janik, Sayigh, & Wells, 2006; Wanker, Sugama, & Prinage, 2005) – a unique signal that may be used to identify individuals – the signature whistle (Caldwell & Caldwell, 1965).

In the years since they were first identified, it has become relatively well established that signature whistles are utilized in much the same way as contact calls, except they also carry identity information (Janik et al., 2006). One study even showed that when separate groups of dolphins encountered each other in the wild, an increase in signature whistling was correlated with the likelihood that those groups would then travel together (Quick & Janik, 2012). This pattern suggested that signature whistles may be used as greeting communications, potentially soothing fraught initial interactions.

Our study questioned whether this pattern of behavior would be observed in a captive environment. Because dolphin introductions in captive environments occur in a controlled setting, researchers can monitor the production of signature whistles by individuals in this context. We recorded vocal interactions as a new individual, a juvenile male, was introduced to a group of two resident dolphins,

an adult male and female. As observed in the wild, we expected to record an increase in signature whistling at the onset of the introduction.

Instead, we found more individualistic reactions. The new individual, who had been signature whistling constantly during his quarantine, ceased signature whistling immediately upon introduction (X^2 (df=3, N = 475) = 162.095, p < .001, Cohen's w = .58) and the adult male' production of signature whistles did not change. The adult female was the only dolphin to meet our expectation of increased signature whistle production (X^2 (df=3, N = 26) = 16.15, p = .001, Cohen's w = .79) (Table 1).

	Adult	Adult	New	Other
	Female's SW	Male's SW	Dolphin's SW	Whistles
Baseline	0	3	173	46
Introduction Day	11	5	2	32
Post-Introduction	11	3	57	55
Follow-up	10	8	282	85
Total	33	19	514	218

Table 1. Whistle rates per hour for each dolphins' signature whistle (SW) and all other whistles during each phase of the study.

Several possibilities exist for this pattern of whistling. It is possible that the artificial nature of the introduction changed the dolphins' responses. However, more likely, the newcomer's youth and propensity to whistle less when under stress led to his reaction. The adult male's underwhelming response to the new dolphin was unexpected in that most male-male introductions that have been described entail aggressive responses. It is possible that, due to difficulties in detection, the adult male's whistle rate may have been underestimated, however, he also failed to show interest behaviorally, suggesting a true lack of interest. The adult female, consistent with her increased signature whistle production, interacted with the newcomer almost immediately, showing the typical response of dolphins encountering a new conspecific and supporting the hypothesis that signature whistles may be used as a greeting behavior.

This reaction to an introduction is an intriguing glimpse into the function of signature whistles during initial encounters. Future work should allow researchers to continue to delineate the function and use of this unique form of communication. As one of the only species to utilize labels akin to human names, this may elucidate the evolutionary path to flexible, meaningful, referential communication systems.

- Balsby, T. J. S., & Bradbury, J. W. (2009). Vocal matching by orange-fronted conures (Aratinga canicularis). *Behavioural Processes*, *82*(2), 133–9. https://doi.org/10.1016/j.beproc.2009.05.005
- Caldwell, M. C., & Caldwell, D. K. (1965). Individualized Whistle Contours in Bottle-nosed Dolphins (Tursiops truncatus). *Nature*, 207(4995), 434–435. https://doi.org/10.1038/207434a0
- Gregg, J. D. (2013). Are Dolphins Really Smart? Oxford, UK: Oxford University Press.
- Herman, L. M. (1987). Receptive competencies of language-trained animals. *Advances in the Study of Behavior*, *17*, 1–60.
- Herman, L. M., Kuczaj, S. A., & Holder, M. D. (1993). Responses to anomalous gestural sequences by a language-trained dolphin: evidence for processing of semantic relations and syntactic information. *Journal of Experimental Psychology: General*, 122(2), 184–194.
- Janik, V. M., Sayigh, L. S., & Wells, R. S. (2006). Signature whistle shape conveys identity information to bottlenose dolphins. *Proceedings of the National Acadamy of Sciences*, 103, 8293–8297.
- Kako, E. (1999). Elements of syntax in the systems of three language-trained animals. *Animal Learning and Behavior*, 27(1), 1–14.
- Quick, N. J., & Janik, V. M. (2012). Bottlenose dolphins exchange signature whistles when meeting at sea. *Proceedings of the Royal Society Biological Sciences Series B*, 279(1738), 2539–2545. Retrieved from http://lynx.lib.usm.edu:2048/login?url=http://search.ebscohost.com/login.a spx?direct=true&db=bxh&AN=BACD201200325753&site=ehost-live
- Wanker, R., Sugama, Y., & Prinage, S. (2005). Vocal labelling of family members in spectacled parrotlets, Forpus conspicillatus. *Animal Behaviour*, 70(1), 111–118. https://doi.org/10.1016/j.anbehav.2004.09.022

NEUROIMAGING STUDIES ON DRAWING SUGGEST A LINK BETWEEN SYMBOLIC BEHAVIOR TO LANGUAGE: A POTENTIAL HINT TO THE NEANDERTHAL LANGUAGE CONTROVERSY

Michiru Makuuchi*

*Corresponding Author: makuuchi-michiru@rehab.go.jp Neuropsychology, National Rehabilitation Center for Persons with Disabilities, Tokorozawa, Japan

The absence of symbolic behaviors has often been associated with the lack of language in the Neanderthals. However, the conceptual link between symbolic behavior and language is simply speculative, therefore it calls for more convincing arguments (Balari, et al., 2011; Berwick, et al., 2013). Regarding the differences in the symbolic behaviors in the two species, drawing shows a striking contrast probably because its privileged position for preservation compared to other symbolic behaviours such as gesture, music, dance, and ritual. The homo sapiens have demonstrated their competence for symbolic behavior by depicting figurative pictures such as humans and animals on cave walls in Europe and Indonesia as early as 40 kyr ago (Aubert, et al., 2014). In contrast, clear archeological records for the Neanderthal artefacts is extremely scanty (Higham, et al., 2010), except stone and bone tools (Soressi, et al., 2013) (but see (Hoffmann, et al., 2018; Jaubert, et al., 2016)). This discrepancy becomes more impressive when we take into account of their temporal overlap estimated 2,600-5,400 yrs (Higham, et al., 2014) or of their direct and perhaps intimate interaction implied by the gene flow from the Neanderthals to the modern humans (Prufer, et al., 2014). The absence of evidence is not the evidence of absence, and drawing might be done on sand or on other frail materials, but the lack of drawing works in Neanderthal's sites is notable when we consider discovery of thousands of their fossils. How do cognitive abilities that create figurative art relate to language faculty? Recent neuroimaging studies on drawing may hold a useful clue to this question.

274

Neuroimaging is a technique to measure brain activities during various perceptual or cognitive tasks. Functional magnetic resonance imaging (fMRI) is the most popular method among others since it allows the best spatial resolution in localizing activated loci in the brain without any invasion to the body. Several studies on the brain mechanisms of drawing have revealed that drawing activates a network of brain regions. The first study aimed at examining the laterality of activation in the parietal lobes which was the main regions for drawing deficits when the brain is insulted (De Renzi, 1997; Kleist, 1934). Unexpectedly, the authors found activation in the left ventral premotor area (Brodmann area 44, the posterior part of Broca's area) and the right (and to a lesser degree in the left) posterior temporal gyrus in addition to the bilateral parietal activation (Makuuchi, et al., 2003). These two regions were repeatedly reveled co-activated in the subsequent fMRI studies on drawing (Farias, et al., 2006; Harrington, et al., 2009; Miall, et al., 2009; Schaer, et al., 2012; Yuan and Brown, 2015). Because of their anatomical locations, we reason they are connected by the arcuate fasciculus which subserves language (Catani, et al., 2007). Here we obtain two insights. The first is that acquisition of language and drawing skill in the modern humans might result from the evolution of the arcuate fasciculus (Makuuchi, 2010; Rilling, et al., 2008). The second is that language and drawing may share similar fundamental computation in the brain, for instance building hierarchical structures of elements (e.g. words/strokes) for the externalization of mental representations by effectors (e.g. speech organs/hands) (Makuuchi, et al., 2003). These insights can be directly tested by neuroimaging methods. Researchers may be able to formulate hierarchical structure in drawing and examine if the processing of hierarchical structure in drawing activates Broca's area as in sentence processing (Friederici, et al., 2006; Makuuchi, et al., 2009; Zaccarella, et al., 2017) and if the activated regions in Broca's area have anatomical connections to the co-activated posterior regions.

- Aubert, M., Brumm, A., Ramli, M., Sutikna, T., Saptomo, E.W., Hakim, B., Morwood, M.J., van den Bergh, G.D., Kinsley, L., Dosseto, A. (2014) Pleistocene cave art from Sulawesi, Indonesia. Nature, 514:223-7.
- Balari, S., Benitez-Burraco, A., Camps, M., Longa, V.M., Lorenzo, G., Uriagereka, J. (2011) The archaeological record speaks: bridging anthropology and linguistics. International journal of evolutionary biology, 2011:382679.

- Berwick, R.C., Hauser, M.D., Tattersall, I. (2013) Neanderthal language? Justso stories take center stage. Frontiers in psychology, 4:671.
- Catani, M., Allin, M.P., Husain, M., Pugliese, L., Mesulam, M.M., Murray, R.M., Jones, D.K. (2007) Symmetries in human brain language pathways correlate with verbal recall. Proceedings of the National Academy of Sciences of the United States of America, 104:17163-8.
- De Renzi, E. (1997) Visuospatial and constructional disorders. In: Feinberg, T., Farah, M., editors. Behavioral Neurology and Neuropsychology. New York: McGraw-Hill. p 297–307.
- Farias, D., Davis, C., Harrington, G. (2006) Drawing: its contribution to naming in aphasia. Brain and language, 97:53-63.
- Friederici, A.D., Bahlmann, J., Heim, S., Schubotz, R.I., Anwander, A. (2006) The brain differentiates human and non-human grammars: functional localization and structural connectivity. Proceedings of the National Academy of Sciences of the United States of America, 103:2458-63.
- Harrington, G.S., Farias, D., Davis, C.H. (2009) The neural basis for simulated drawing and the semantic implications. Cortex; a journal devoted to the study of the nervous system and behavior, 45:386-93.
- Higham, T., Douka, K., Wood, R., Ramsey, C.B., Brock, F., Basell, L., Camps, M., Arrizabalaga, A., Baena, J., Barroso-Ruiz, C., Bergman, C., Boitard, C., Boscato, P., Caparros, M., Conard, N.J., Draily, C., Froment, A., Galvan, B., Gambassini, P., Garcia-Moreno, A., Grimaldi, S., Haesaerts, P., Holt, B., Iriarte-Chiapusso, M.J., Jelinek, A., Jorda Pardo, J.F., Maillo-Fernandez, J.M., Marom, A., Maroto, J., Menendez, M., Metz, L., Morin, E., Moroni, A., Negrino, F., Panagopoulou, E., Peresani, M., Pirson, S., de la Rasilla, M., Riel-Salvatore, J., Ronchitelli, A., Santamaria, D., Semal, P., Slimak, L., Soler, J., Soler, N., Villaluenga, A., Pinhasi, R., Jacobi, R. (2014) The timing and spatiotemporal patterning of Neanderthal disappearance. Nature, 512:306-9.
- Higham, T., Jacobi, R., Julien, M., David, F., Basell, L., Wood, R., Davies, W., Ramsey, C.B. (2010) Chronology of the Grotte du Renne (France) and implications for the context of ornaments and human remains within the Chatelperronian. Proceedings of the National Academy of Sciences of the United States of America, 107:20234-9.
- Hoffmann, D.L., Standish, C.D., Garcia-Diez, M., Pettitt, P.B., Milton, J.A., Zilhao, J., Alcolea-Gonzalez, J.J., Cantalejo-Duarte, P., Collado, H., de Balbin, R., Lorblanchet, M., Ramos-Munoz, J., Weniger, G.C., Pike, A.W.G. (2018) U-Th dating of carbonate crusts reveals Neandertal origin of Iberian cave art. Science (New York, N.Y.), 359:912-915.
- Jaubert, J., Verheyden, S., Genty, D., Soulier, M., Cheng, H., Blamart, D., Burlet, C., Camus, H., Delaby, S., Deldicque, D., Edwards, R.L., Ferrier, C., Lacrampe-Cuyaubere, F., Leveque, F., Maksud, F., Mora,

P., Muth, X., Regnier, E., Rouzaud, J.N., Santos, F. (2016) Early Neanderthal constructions deep in Bruniquel Cave in southwestern France. Nature, 534:111-4.

Kleist, K. (1934) Gehirnpathologie. Leipzig. Barth.

- Makuuchi, M. (2010) fMRI studies on drawing revealed two new neural correlates that coincide with the language network. Cortex; a journal devoted to the study of the nervous system and behavior, 46:268-9.
- Makuuchi, M., Bahlmann, J., Anwander, A., Friederici, A.D. (2009) Segregating the core computational faculty of human language from working memory. Proceedings of the National Academy of Sciences of the United States of America, 106:8362-7.
- Makuuchi, M., Kaminaga, T., Sugishita, M. (2003) Both parietal lobes are involved in drawing: a functional MRI study and implications for constructional apraxia. Brain research. Cognitive brain research, 16:338-47.
- Miall, R.C., Gowen, E., Tchalenko, J. (2009) Drawing cartoon faces--a functional imaging study of the cognitive neuroscience of drawing. Cortex; a journal devoted to the study of the nervous system and behavior, 45:394-406.
- Prufer, K., Racimo, F., Patterson, N., Jay, F., Sankararaman, S., Sawyer, S., Heinze, A., Renaud, G., Sudmant, P.H., de Filippo, C., Li, H., Mallick, S., Dannemann, M., Fu, Q., Kircher, M., Kuhlwilm, M., Lachmann, M., Meyer, M., Ongyerth, M., Siebauer, M., Theunert, C., Tandon, A., Moorjani, P., Pickrell, J., Mullikin, J.C., Vohr, S.H., Green, R.E., Hellmann, I., Johnson, P.L., Blanche, H., Cann, H., Kitzman, J.O., Shendure, J., Eichler, E.E., Lein, E.S., Bakken, T.E., Golovanova, L.V., Doronichev, V.B., Shunkov, M.V., Derevianko, A.P., Viola, B., Slatkin, M., Reich, D., Kelso, J., Paabo, S. (2014) The complete genome sequence of a Neanderthal from the Altai Mountains. Nature, 505:43-9.
- Rilling, J.K., Glasser, M.F., Preuss, T.M., Ma, X., Zhao, T., Hu, X., Behrens, T.E. (2008) The evolution of the arcuate fasciculus revealed with comparative DTI. Nature neuroscience, 11:426-8.
- Schaer, K., Jahn, G., Lotze, M. (2012) fMRI-activation during drawing a naturalistic or sketchy portrait. Behavioural brain research, 233:209-16.
- Soressi, M., McPherron, S.P., Lenoir, M., Dogandzic, T., Goldberg, P., Jacobs, Z., Maigrot, Y., Martisius, N.L., Miller, C.E., Rendu, W., Richards, M., Skinner, M.M., Steele, T.E., Talamo, S., Texier, J.P. (2013) Neandertals made the first specialized bone tools in Europe. Proceedings of the National Academy of Sciences of the United States of America, 110:14186-90.
- Yuan, Y., Brown, S. (2015) Drawing and writing: An ALE meta-analysis of sensorimotor activations. Brain and cognition, 98:15-26.

Zaccarella, E., Meyer, L., Makuuchi, M., Friederici, A.D. (2017) Building by Syntax: The Neural Basis of Minimal Linguistic Structures. Cerebral cortex (New York, N.Y. : 1991), 27:411-421.

WHAT CAN SOUND SYMBOLISM TELL US ABOUT THE EVOLUTION OF LANGUAGE? A COMPARATIVE STUDY IN HUMANS AND GREAT APES

KONSTANTINA MARGIOTOUDI^{*1,2}, MANUEL BOHN⁴, MATTHIAS ALLRITZ⁵ and FRIEDEMANN PULVERMÜLLER^{1,2,3}

*Corresponding Author: konstantina.margiotoudi@fu-berlin.de
¹Brain Language Laboratory, Department of Philosophy and Humanities, Frei Universität,Berlin,Germany
²Berlin School of Mind and Brain, Humboldt Universität, Berlin, Germany
³Einstein Center for Neurosciences,Berlin, Germany
⁴Leipziger Foschungszentrum für frühkindliche Entwicklung, Universität Leipzig, Germany
⁵School of Psychology & Neuroscience, University of St Andrews, St Andrews, Scotland

There has been a long debate in the literature of semantics between the arbitrariness of language proposed by Saussure (1959) and the non-arbitrary associations, known as sound symbolism. Sound symbolism is a form of iconicity based on similarity between a linguistic form and the sensory-motor properties of its referent (Perniss & Vigliocco, 2014). A classic example of sound symbolism is the sound-shape correspondence, as described by Köhler (1929), where the non-word "maluma" was judged to be a good match with a round shape whereas the non-word "takete" matched better to a spiky shape.

Recent theories have highlighted the relevant role of sound symbolism and that of iconicity for the evolution of language (Ramachandran & Hubbard, 2001; Arbib, 2005; Imai & Kita, 2014; Perniss & Vigliocco 2014).

In other words, iconic signals are proposed as the starting point for the evolution of language. Sound symbolic mimicking of the external word with the movement of lips and tongue is another form of iconicity that would fit these theories.

Moreover linked to theories on embodied cognition (Barsalou et al., 2003), language processing is proposed to be based on distributed neural circuits grounded in the action and perception system of the brain and mind (Pulvermüller, 2018). Sound symbolism has been proposed as the linkage between language and human sensory-motor experience, playing a determinant role on the evolution of language (Perniss & Vigliocco, 2014).

Despite this theoretical interest on the phylogenetic origins of sound

symbolism, as far as we are concerned there have been no comparative studies in testing sound symbolic associations in humans and great apes. In addition, considering the new evidence on the perception of iconic gestures in chimpanzees (Bohn et al., 2016), the present study aims to explore whether our closest living relatives are able of perceiving vocal iconic mappings between shapes and sounds.

For that purpose, we ran a two-alternative forced choice (2AFC) audiovisual task, the classic "maluma"-"takete" paradigm (Köhler, 1929). 24 healthy human subjects, four chimpanzees (Pan troglodytes) and two gorillas (Gorilla gorilla gorilla) took part in the study. During the task the subjects listened to a word copresented with two shapes, one edgy and one round, and they had to choose one of the two co-presented shapes. In order to avoid any habituation effect, we generated disyllabic non-words that sounded more "round" or "edgy" based on combination of vowels and consonants that sound more "round" or "edgy" in accordance with the literature (Nielsen & Rendall, 2013; McCormick et al., 2015).

We also generated shapes that looked edgy or round. Both sounds and shapes were rated with an online questionnaire in order to select those that have the higher degree of "edginess" or "roundness". The human subjects were tested in a behavioral booth by responding with a button box and choosing one of the two shapes that matched the aurally simultaneously presented non-word.

In accordance with previous studies, humans preferred to associate round shapes to words that consisted of round vowels, voiced plosives and sonorants. In contrast, disyllabic words that consisted of edgy vowels and voiceless plosives or fricatives were associated more with edgy shapes.

Based on the findings from humans we are currently conducting an analogous study in great apes at the Wolfgang Kohler Primate Research Center, Zoo Leipzig, Germany. Specifically six chimpanzees (Pan troglodytes, 3 males) and two gorillas (Gorilla gorilla gorilla, 2 females) that were touch screen trained, are tested with the same stimuli in a comparable task. We modified the initial task by increasing the number of trials, in order to have data comparable to the human ones, and by rewarding with food 50 % of the trials in order to keep them engaged to the task.

Based on the results of this study, we will shed light to different hypothesis regarding the role of iconicity and that of sound symbolism in the phylogenetic origins of language. Moreover, new questions will raise regarding the mechanisms that boost the evolution of language. Perhaps the different characteristics of the articulatory system (Fitch, 2000) or differences in the

brain's neuroanatomical structure for spoken language and verbal working memory between human and non human primates (Schomers & Pulvermüller, 2016) may have limited the role of verbal iconic communication in the evolution of language, so that apes would not show evidence of sound symbolic processing.

- Arbib, M. A. (2005). From monkey-like action recognition to human language An evolutionary framework for neurolinguistics. *Behavioral and brain sciences*, 28(2), 105-124.
- Barsalou, L. W., Simmons, W. K., Barbey, A. K., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in cognitive sciences*, 7(2), 84-91.
- Bohn, M., Call, J., & Tomasello, M. (2016). Comprehension of iconic gestures by chimpanzees and human children. *Journal of experimental child psychology*, 142, 1-17.
- Fitch, W. T. (2000). The evolution of speech: A comparative review. *Trends in Cognitive Sciences*, 4(7), 258–267.
- Imai, M., & Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Phil. Trans. R. Soc. B*, 369(1651), 20130298.
- Köhler, W. (1929). Gestalt psychology. New York, USA: Liveright
- McCormick, K., Kim, J., List, S., & Nygaard, L. C. (2015). Sound to Meaning Mappings in the Bouba-Kiki Effect. In *CogSci*.
- Nielsen, A. K., & Rendall, D. (2013). Parsing the role of consonants versus vowels in the classic Takete-Maluma phenomenon. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 67(2), 153.
- Perniss, P., & Vigliocco, G. (2014). The bridge of iconicity: from a world of experience to the experience of language. *Phil. Trans. R. Soc. B*, 369(1651), 20130300.
- Pulvermüller, F. (2017). Neural Reuse Of Action Perception Circuits For Language, Concepts And Communication. *Progress in Neurobiology*.
- Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia--a window into perception, thought and language. *Journal of consciousness studies*, 8(12), 3-34.
- Saussure, F. D. (1959). Course in General Linguistics. Ed. Charles Bally, Albert Sechehaye, and Albert Riedlinger. Trans. Wade Baskin. New York: Philosophical Library.
- Schomers, M. R., & Pulvermüller, F. (2016). Is the sensorimotor cortex relevant for speech perception and understanding? An integrative review. *Frontiers in human neuroscience*, 10.

CULTURAL TRANSMISSION OF MELODIC AND RHYTHMIC UNIVERSALS: 4 EXPERIMENTS AND A MODEL

TANIA DELGADO¹, ANDREA RAVIGNANI^{*2,3,4}, TESSA VERHOEF^{#5,6}, BILL THOMPSON^{2,4}, THOMAS GROSSI⁷, SIMON KIRBY⁷

*Corresponding Author: andrea.ravignani@gmail.com #Corresponding Author: tessauva@gmail.com ¹Department of Cognitive Science, University of California, San Diego, La Jolla, USA

²Artificial Intelligence Lab, Vrije Universiteit Brussel, Brussels, Belgium ³Research Department, Sealcentre Pieterburen, Pieterburen, Netherlands

⁴Language and Cognition Department, Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands

⁵Leiden Institute of Advanced Computer Science (LIACS), Leiden University, Leiden, The Netherlands

⁶Center for Research in Language, University of California, San Diego, La Jolla, USA ⁷Centre for Language Evolution, School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, UK.

Humans are well-versed at processing sequences. In the lab, when confronted with language or other culturally-transmitted systems, humans introduce and amplify structural regularities making the systems easier to learn (Kirby, Griffiths, & Smith, 2014; Kirby, 2017). Is the cultural emergence of fine-grained regularities a prerogative of language alone? Can cultural transmission explain universals in musical structure (Savage et al., 2015; Trehub, 2015) as it explains 'linguistic universals' (in the Greenbergian sense)? We tackled these questions in the lab and in-silico by adopting an iterated learning paradigm. Two experiments addressed the evolution of *rhythmic* structure (Ravignani et al., 2017), and two the evolution of *melodic* structure (using, among others, data from Verhoef, 2012; Verhoef et al., 2014). Depending on the experiment, participants were given a slide whistle or an electronic drum kit, and were asked to imitate a sound sequence to the best of their abilities. The output of one generation of participants became the input of the next generation.

Across experiments we varied conditions known to affect transmission chains, such as immediate vs. delayed recall, within-participant vs. between-participant transmission, etc. We also introduce a probabilistic model for the latent structures underpinning rhythmic sequences, alongside a psychologically plausible posterior sampling algorithm (Ravignani et al., 2017). This allows us to obtain approximate structural descriptions of rhythmic patterns across conditions and generations. The initially random experimental stimuli became more musical. In particular, most regularities emerging from our transmission chains are statistical universals of world music (Savage et al., 2015; Fitch, 2017). Both whistles and drumming patterns became more compressible, measured in terms of decreasing entropy, and easier to learn. We found reuse of a small set of basic building blocks in the emerging systems, resulting in more predictable sequences of sounds (Ravignani, 2017). This corresponds with a musical universal, namely the repetition of melodic and rhythmic phrases in music. This reuse of elements is accompanied by a transition from continuous to discrete use of pitch contours (showing convergence to another universal; Ravignani & Verhoef, 2017). Participants produced sequences containing melodic and rhythmic patterns, i.e. musical motifs. We also see a gradual increase in mirrored elements, which suggests that the emerging melodies contain arched contours (a common musical universal). In addition, drumming sequences became more isochronous (Ravignani & Madison, 2017), and composed of few (categorically distributed) alternating inter-beat intervals, related by small integer ratios. Patterns transformed by between-participants transmission show similar properties to those emerging from within-participant transmission (i.e. self learning; Ravignani et al., 2017). Other melodic universals, such as the length of phrases, frequency intervals of melody contours, and organization of scales can also be measured in this data set and are currently being tested. Analysis of the probabilistic model supports these insights, suggesting that later generations show increased re-use of prototypical building blocks both within and across individual sequences (Ravignani et al., 2017). The emergence of musical structure via cultural transmission: (a) does not require semantics or learning language-like behaviours; (b) operates similarly across domains of human cognition; (c) explains characteristics of music appearing as statistical universals around the world (Savage et al., 2015).

Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant

agreement No. 665501 with the research Foundation Flanders (FWO) (Pegasus2 Marie Curie fellowship 12N5517N awarded to A.R.), a visiting fellowship in Language Evolution from the Max Planck Society (awarded to A.R.), and ERC Grant [283435] ABACUS (awarded to B.d.B.).

- Fitch, W. T. (2017). Cultural evolution: Lab-cultured musical universals. *Nature Human Behaviour*, *1*, 0018.
- Kirby, S., Griffiths, T., & Smith, K. (2014). Iterated learning and the evolution of language. *Current opinion in neurobiology*, 28, 108-114.
- Kirby, S. (2017). Culture and biology in the origins of linguistic structure. *Psychonomic bulletin & review*, 24(1), 118-137.
- Ravignani, A., Delgado, T., & Kirby, S. (2016). Musical evolution in the lab exhibits rhythmic universals. *Nature Human Behaviour*, *1*, 0007.
- Ravignani, A. (2017). Visualizing and interpreting rhythmic patterns using phase space plots. *Music Perception: An Interdisciplinary Journal*, 34(5), 557-568.
- Ravignani, A., Thompson, B., Grossi, T., Delgado, T., & Kirby, S. (2017). Evolving building blocks of rhythm: How human cognition creates music via cultural transmission. *bioRxiv*, 198390.
- Ravignani, A., & Madison, G. (2017). The paradox of isochrony in the evolution of human rhythm. *Frontiers in Psychology*, 8.
- Ravignani, A., & Verhoef, T. (2017). Which melodic universals emerge from repeated signaling games? (No. e3366v1). *PeerJ Preprints*.
- Savage, P. E., Brown, S., Sakai, E., & Currie, T. E. (2015). Statistical universals reveal the structures and functions of human music. *Proceedings of the National Academy of Sciences*, 112(29), 8987-8992.
- Trehub, S. E. (2015). Cross-cultural convergence of musical features. Proceedings of the National Academy of Sciences, 112(29), 8809-8810.
- Verhoef, T. (2012) The origins of duality of patterning in artificial whistled languages. *Language and Cognition* 4(4), 357-380.
- Verhoef, T., Kirby, S. & de Boer, B. (2014) Emergence of combinatorial structure and economy through iterated learning with continuous acoustic signals. *Journal of Phonetics* 43C, pp. 57-68

SOUND PRODUCTION LEARNING AND THE CONTIGUUM HYPOTHESIS

PEDRO TIAGO MARTINS*1,2 and CEDRIC BOECKX1,2,3

*Corresponding Author: pmartima23@alumnes.ub.edu
¹Section of General Linguistics, University of Barcelona, Barcelona, Spain
²University of Barcelona Institute of Complex Systems, Barcelona, Spain
³ICREA, Barcelona, Spain

1. Introduction

Besides humans, vocal learning is also attested in some bird lineages (songbirds, parrots, hummingbirds), some cetaceans, bats, pinnipeds and some elephants (Ravignani et al., 2016). Vocal learning is usually seen as a transparent behavioral trait, with an associated neural substrate, and typically a species is said to either have it or not (i.e. species are said to be "vocal learners" or "vocal non-learners"). This classification guides research on a variety of topics when it comes to understanding the evolution of vocal learning and its relationship with other language-related traits. However, there have been attempts at a more nuanced view, resulting in non-dichotomous typologies of vocal learning that include more species and a wider spectrum of capacities. A notable example is the *continuum hypothesis* put forward by Arriaga and Jarvis (2013), for which they propose a more nuanced scheme of vocal learning, going beyond the traditional all-or-nothing view and incorporating cases of species which do not conform to that reductionist classification, namely those who can produce novel vocalizations without mimicry (see Petkov and Jarvis (2012) for a review of evidence in this direction).

2. Proposal

In this work, we continue in the same vein and recontextualize vocal learning as a case of *sound production learning*, a more general sound production ability that does not necessarily rest on vocal control. We present this as part of a *contiguum hypothesis*, as a way of emphasizing abilities that border on and interact with others by virtue of sharing common ground at different levels. This will allow us to bring our close relatives into the fold, and foster non-human primate research on aspects (behavioral, ecological, neural, and genetic) that play an important role in language and allow us to peer into its evolution in more permissive ways (see Lameira, 2017 for related discussion and literature). There are several reasons for attempting to do so.

1. It has been shown in recent years that the vocal tract of many non-human primates is not a obstacle (Fitch, de Boer, Mathur, & Ghazanfar, 2016) to vocal abilities. If vocal anatomy is abandoned as a proxy for species worth studying, the number of interesting species for the study of language evolution increases tremendously.

2. Another important reason has to do with volition. For example, even though many primates lack the laryngeal control necessary for bona fide vocal learning, some can still modulate calls by external means, such as using their hands or a leaf in front of their mouth. This contributes e.g. to size exaggeration (Lameira, Hardus, & Wich, 2011). Orangutans also seem to be able to spontaneously acquire and modulate human-like whistles (Wich et al., 2009). Other species, like gibbons, are able to propagate their vocalizations across long distances by reconfiguring their vocal tract in ways akin to soprano singing (Koda et al., 2012). We will present examples showing that some species have ways of going around their lack of laryngeal control by other means and still produce sounds beyond their innate repertoire. We will propose that *vocal* control is not the only piece of the puzzle of controlled vocalizations.

3. One other, crucial reason has to do with the brain. The brain "signature" of vocal learning species is a direct connection between the motor cortex and the larynx (in birds, the arcopallium and syrinx), which endows them with the necessary vocal control. Opening up the set of interesting abilities to a wider *sound production learning* capacity invites the exploration of different brain structures and pathways putatively relevant.

3. Final remarks

As with many other dichotomies (e.g. innate/learned, nature/nurture), a strict, allor-nothing classification of vocal learning misses important distinctions that rest on finer-grained behavioral and neuroanatomical traits, something which in turn Arriaga and Jarvis (2013) capture with their *continuum hypothesis*. We extend this way of proceeding beyond specific traits, and generalize it, using a generic notion of vocal abilities as proof of concept.

By examining cognition as a "contiguous" space, where borders may be apparent at any one level, but vanish when others are taken into account, we intend to capture the non-hierarchical, multi-directional character of the evolution of complex traits. Instead of thinking of straight lines from stage A to stage B, or trees that bifurcate at node A and yield A and B, a more fruitful approach when studying complex traits is to think of a territory, where dividing lines do exist, but traversing them is the rule rather than the exception. We find this particularly useful in the study of language evolution, which clearly defies the "speciality" that known dichotomies (e.g. Faculty of Language in the Narrow Sense (FLN)/Faculty of Language in the Broad Sense (FLB), Hauser et al. 2002) so vehemently embody, with little progress.

- Arriaga, G., & Jarvis, E. D. (2013). Mouse vocal communication system: Are ultrasounds learned or innate? *Brain and language*, 124(1). doi: 10.1016/ j.bandl.2012.10.002
- Fitch, W. T., de Boer, B., Mathur, N., & Ghazanfar, A. A. (2016). Monkey vocal tracts are speech-ready. *Science Advances*, 2(12), e1600723. doi: 10.1126/sciadv.1600723
- Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science*, *298*, 1569–1579.
- Koda, H., Nishimura, T., Tokuda, I. T., Oyakawa, C., Nihonmatsu, T., & Masataka, N. (2012). Soprano singing in gibbons. *American Journal of Physical Anthropology*, 149(3), 347–355. doi: 10.1002/ajpa.22124
- Lameira, A. R. (2017). Bidding evidence for primate vocal learning and the cultural substrates for speech evolution. *Neuroscience & Biobehavioral Reviews*, 83(Supplement C), 429–439. doi: 10.1016/j.neubiorev.2017.09.021
- Lameira, A. R., Hardus, M. E., & Wich, S. A. (2011). Orangutan Instrumental Gesture-Calls: Reconciling Acoustic and Gestural Speech Evolution Models. *Evolutionary Biology*, 39(3), 415–418. doi: 10.1007/ s11692-011-9151-6
- Petkov, C. I., & Jarvis, E. (2012). Birds, primates, and spoken language origins: Behavioral phenotypes and neurobiological substrates. *Frontiers in Evolutionary Neuroscience*, 4. doi: 10.3389/fnevo.2012.00012
- Ravignani, A., Fitch, W. T., Hanke, F. D., Heinrich, T., Hurgitsch, B., Kotz, S. A.,
 ... de Boer, B. (2016). What Pinnipeds Have to Say about Human Speech,
 Music, and the Evolution of Rhythm. *Frontiers in Neuroscience*, 10. doi: 10.3389/fnins.2016.00274
- Wich, S. A., Swartz, K. B., Hardus, M. E., Lameira, A. R., Stromberg, E., & Shumaker, R. W. (2009). A case of spontaneous acquisition of a human sound by an orangutan. *Primates*, 50(1), 56–64. doi: 10.1007/s10329-008 -0117-y

THE HUMAN ARCUATE FASCICULUS PROVIDES SPECIFIC ADVANTAGES TO PROCESS COMPLEX SEQUENTIAL STIMULI, NOT HIERARCHIES IN GENERAL

MAURICIO DIAS MARTINS^{*1,2}, ARNO VILLRINGER^{1,2}

*Corresponding Author: mmartins@cbs.mpg.de

¹Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany ²Berlin School of Mind and Brain, Humboldt Universität zu Berlin, Berlin, Germany

Hierarchies are sets or sequences of elements connected in the form of a rooted tree. They possess the key properties: (1) all elements are combined into one structure; (2) one element is superior to all others; and (3) no element is superior to itself (that is, there are no cycles, direct or indirect)" (Fitch & Martins, 2014). Defined as such, hierarchies exist in multiple domains. Linguistic syntax, and tonal and action sequences display a multi-layered set-of-sets organization. Moreover, social (e.g. family and company structures) and spatial hierarchies (e.g. landmark-based navigation) also display asymmetrical and multi-layered relations between different elements and sets of elements.

Humans can represent the hierarchical structure in all these domains, and to extend their hierarchical depth when necessary. In the same way that we can extend any arbitrarily long sentence, we can also join any two arbitrarily complex social groups such as the armies of two countries to form a joint inter-national army (or inter-continental, inter-planetary, inter-galactic, etc.).

Humans are especially capable of generating hierarchies. While we are able to assemble these kinds of structures in language, music and complex action (Fitch & Martins, 2014), analogous capacities are missing in other species (Fitch & Friederici, 2012), even though they can process simpler structures to some extent (Wilson, Marslen-Wilson, & Petkov, 2017).

The cognitive and neural substrata supporting this capacity are a matter of active research and discussion. In neurolinguistics, this capacity is usually mapped to the ventral portions of Brodmann's area 44 (BA44), and its interactions with the posterior Superior Temporal Sulcus (Fitch, 2017; Friederici, 2017; Milne et al., 2016). Interestingly, these two regions are connected by a fiber tract, called the Arcuate Fasciculus (AF), which is exceptionally well-developed in humans (Rilling et al., 2008).

The available data suggests the hypothesis that the human ability to represent linguistic hierarchy evolved over a general sequence-processing machinery already available in the primate brain, to which a highly-developed AF was added (Wilson et al., 2017). Some extended this framework to music and action, where hierarchical processing also recruits regions within the Inferior Frontal Gyrus (Fadiga, Craighero, & D'Ausilio, 2009; Fitch & Martins, 2014).

Here, we present a critical challenge to this hypothesis. Consider that there are two groups of domains in which humans can represent hierarchies. In the first, signals are composed of ordered sequences. Here, the serial order of the physical stimuli determines the perceived content or meaning ('Mary likes John' vs. 'John Mary likes'). Even though linguistic hierarchies are not serial themselves, the signal through which they are communicated and decoded is. In the second group, the presentation order of the elements within the set does not necessarily determine the final structure (think of visual or spatial landscapes, or social structures). While the exact serial input order is crucial to determine the structure of ordered sequences, the same is not true for other hierarchical sets.

This taxonomy is important because while BA44 and the AF seem important to process hierarchies within the first group, they are mostly absent in the second (Kumaran, Melo, & Düzel, 2012; Ligneul, Obeso, Ruff, & Dreher, 2016; Martins et al., 2014). The human ability to represent hierarchies in the visual, spatial and social domains is not supported by these mechanisms but rather by the hippocampus, medial Prefrontal cortex, and other structures. The same has been demonstrated for semantic hierarchies (Neville, et al, 2017).

Taken together, these observations yield a logical puzzle:

1. Primates have a general system to process nonhierarchical sequences.

2. The emergence of the human BA44 and AF allowed for the capacity to represent hierarchies to evolve in language.

3. The human ability to represent hierarchies in some domains does not activate the brain areas connected via the AF.

There are two ways to solve this puzzle: The first is to assume that the capacity to represent hierarchies evolved several times, once within language, and for other domains in other time periods. The second entails that the capacity to process hierarchies was first present in the visual, spatial and social domains and then specific changes in BA44 and AF made this capacity available for language (or in general for domains hinging on specific serial order of the input).

In either case, BA 44 and AF seem to be important to process complex structured sequences, but not hierarchies in general. On the one hand, this neural system might be involved in the core generative capacity for hierarchical processing, but only in language. On the other hand, it might connect a previously available capacity to represent sets of sets with a robust capacity to parse sequential information. The latter would be especially important when sequences contain hierarchical relations between elements that are distant in the serial order.

- Fadiga, L., Craighero, L., & D'Ausilio, A. (2009). Broca's area in language, action, and music. Annals of the New York Academy of Sciences, 1169(1), 448-458. doi:10.1111/j.1749-6632.2009.04582.x
- Fitch, W.T. (2017). Empirical approaches to the study of language evolution. *Psychon Bull Rev, 24*(1), 3-33. doi:10.3758/s13423-017-1236-5
- Fitch, W.T., & Friederici, A.D. (2012). Artificial grammar learning meets formal language theory: an overview. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1598), 1933-1955. doi:10.1098/rstb.2012.0103
- Fitch, W.T., & Martins, M.D. (2014). Hierarchical processing in music, language, and action: Lashley revisited. Ann N Y Acad Sci, 1316, 87-104. doi:10.1111/nyas.12406
- Friederici, A. D. (2017). Evolution of the neural language network. *Psychon Bull Rev, 24*(1), 41-47. doi:10.3758/s13423-016-1090-x
- Kumaran, D., Melo, H.L., & Düzel, E. (2012). The emergence and representation of knowledge about social and nonsocial hierarchies. *Neuron*, 76(3), 653-666. doi:10.1016/j.neuron.2012.09.035
- Ligneul, R., Obeso, I., Ruff, C. C., & Dreher, J. C. (2016). Dynamical Representation of Dominance Relationships in the Human Rostromedial Prefrontal Cortex. *Curr Biol*, 26(23), 3107-3115. doi:10.1016/j.cub.2016.09.015
- Martins, M.D., Fischmeister, F.P., Puig-Waldmuller, E., Oh, J., Geissler, A., Robinson, S., Fitch, W.T., Beisteiner, R. (2014). Fractal image perception provides novel insights into hierarchical cognition. *Neuroimage*, 96, 300-308. doi:10.1016/j.neuroimage.2014.03.064
- Milne, A. E., Mueller, J. L., Mannel, C., Attaheri, A., Friederici, A. D., & Petkov, C. I. (2016). Evolutionary origins of non-adjacent sequence processing in primate brain potentials. *Sci Rep, 6*, 36259. doi:10.1038/srep36259
- Neville, D, Theves, S, Fernández, G, & Doeller, C. (2017). *Tracking the emergence of hierarchical conceptual knowledge*. Paper presented at the 23rd Annual Meeting of the Organization for Human Brain Mapping, Vancouver Canada.
- Rilling, J.K., Glasser, M.F., Preuss, T.M., Ma, X., Zhao, T., Hu, X., & Behrens, T.E.J. (2008). The evolution of the arcuate fasciculus revealed with comparative DTI. *Nature Neuroscience*, 11(4), 426-428.
- Wilson, Benjamin, Marslen-Wilson, William D., & Petkov, Christopher I. (2017).Conserved Sequence Processing in Primate Frontal Cortex. Trends in
Neurosciences, 40(2), 72-82.doi:https://doi.org/10.1016/j.tins.2016.11.004

A POTENTIAL LINK BETWEEN INFANT-DIRECTED SPEECH AND EVOLUTION OF VOCAL COMMUNICATION

REIKO MAZUKA^{1*}, TADAHISA KONDO², and AKIKO HAYASHI³

*Corresponding Author: mazuka@brain.riken.jp
¹ Laboratory for language development, RIKEN BSI, Saitama, Japan
² Kogakuin University, Tokyo, Country
³ Tokyo Gakugei University, Tokyo, Japan

Adults modify their speech when they talk to infants and young children. This style of speech is called "Infant-directed speech" (IDS). Research has shown that IDS differs from ADS in multiple ways, for example, phonologically, lexically, syntactically, as well as pragmatically. Many of these features have been argued to contribute to attracting and maintaining infants' attention, communicating affects, as well as facilitating their language acquisition (Soderstrom, 2007, for review). Yet, it is still not clear how adults would know what to say in IDS. In this presentation, we focus on the specific ways IDS favors certain segments over others and ask how adults would know which segments to use selectively in IDS.

The prominent view in the field is that mothers (and other adults) fine-tune their production to infants' capabilities such that segments that are produced early by infants are favored over those that are not produced until later (the "fine-tuning hypothesis," Cross, 1977). The implication of this hypothesis is that mothers' productions are either imitations of children's immature production, or more generally learned through their interaction with their children. In the present paper, we argue for an alternative possibility that is based on Morton's Motivation-Structure Rule Hypothesis, MSRH (1977). He proposed that there is a general relationship between the physical structures of sounds and the motivation underlying their use in animal communication. Considering that human adults' use of IDS is also motivated by their desire to communicate their intensions and affects with infants, the selective use of certain sounds in IDS may be also explained by MSRH. If true, it predicts that IDS would arise from MSRH, and need not be learned from infants.

We tested these predictions in a series of rating studies with Japanese mothers with young infants and college aged Japanese adults with minimal contact with infants and young children. One way in which selective segments are used more frequently in IDS is the use of specialized vocabulary items in IDS, Infant-Directed Vocabulary (IDV). Mothers and college aged adults were asked to rate large sets of nonsense words in terms of how well each item sounds like an IDV on a 7-point Likert scale "sounds like a very good IDV word" (7) to "does not sound like an IDV word at all" (1) (Mazuka, Hayashi, Kondo, 2017a). A different group of college aged adults were asked to rate the same set of words in terms of how good they sound as a Japanese word. Stimuli were created such that we can test whether the following factors would result in higher IDV ratings; 1) presence of specific vowels, 2) consonants, and 3) the prosodic form of the word (Mazuka, Hayashi, Kondo, 2017b).

The results revealed that college aged adults were highly accurate in rating nonsense words as "sounds like a good IDV" when they contained vowels and consonants that are produced early by Japanese children (e.g., /p, b, m/ and /a/), while rating those with late-produced segments (e.g., fricatives (/s, z/) as "does not sound like a good IDV." College aged adults' ratings were highly consistent with the ratings of the mothers as shown in Table 1. The results also revealed that a word that sounds like a good IDV is clearly distinct from those that sound as a good Japanese words.

These results showed that the experience of interacting with children is not necessary for a native speaker of Japanese to have a sense of what a good IDV should sound like in Japanese, which is consistent with the prediction of MSRH. Further explorations into the link between IDS and vocal communication of animals could elucidate the how language communication may have evolved.

	Mothers' IDV-ness	College-age adults' IDV- ness	College-age adults' Japanesenes
Mothers' IDV-ness	1.00	0.867 P<.001	-0.02 n.s.
College-age adults' IDV-ness		1.00	-0.23 p<.001
College-age adults' Japanese-ness			1.00

Table 1. Correlations among Mothers' ratings of IDV-ness, college age adults' IDV-ness, and college age adults' Japanese-ness ratings.

- Cross, T. G. (1977). "Mothers' speech adjustments: The contributions of selected child listener variables." In K. Snow & C. A. Ferguson (eds.), Talking to children: Language input and acquisition (pp. 151–188). Cambridge: Cambridge University Press.Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences*, 13, 707-784.
- Mazuka, R., Kondo, T., & Hayashi, A. (2017a). "Sounds of infant-directed vocabulary: Learned from infants' speech or part of linguistic knowledge?" *Journal of the Phonetic Society of Japan*, 21(1), 45-58.
- Mazuka, R., Kondo, T., & Hayashi, A. (2017b). "Good infant-directed words" do not sound like "good Japanese words." In S. Tanaka et al. (Eds.), On'in kenkyu-no shin-tenkai [New directions in phonological research]. 202–219. Kaitaku-sha, Tokyo.
- Morton, E. S. (1977). "On the occurrence and significance of Motivation-Structural Rules in some bird and mammal songs." *American Naturalist*, 111, 855–869.
- Soderstrom, M. (2007). Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. *Developmental Review*, 27, 501–532.

IDENTIFYING THE APE BEAT IN THE WILD: RHYTHMIC INDIVIDUAL SIGNATURES FROM THE SOUNDS OF MANUAL FRUIT CRACKING IN FONGOLI CHIMPANZEES

ADRIEN MEGUERDITCHIAN*1, ADELYNE VUILLEMIN1, and JILL D. PRUETZ2

*Corresponding Author: adrien.meguerditchian@univ-amu.fr
¹Laboratoire de Psychologie Cognitive, Brain and Language Research Institute, Aix-Marseille Univ, CNRS, Marseille, France
²Department of Anthropology, Texas State University, San Marcos, TX, USA

Rhythm is one of the key components of the temporal structure of language at both production and perception levels. To investigate the phylogenetical origins of this property, comparative approach on the rhythmic properties of vocal communication, acoustic or motor behaviors has been developed between human and a range of animal species including birds, sea lions and nonhuman primates (e.g. Ravignani et al., 2016; Dufour et al., 2015; Cook et al., 2013). In great apes specifically, manual beating - "drumming" - on a tree, object or body part has been documented in different apes' species and contexts including display, play or travels (e.g. Arcadi et al., 1998). "Drumming" in great apes has been considered as a behavior of interest for investigating the ancestral prerequisites to rhythmic processing in humans (Dufour et al., 2015; Fitch 2006). One can question what functional advantage such rhythmic properties may provide to favor the selection of its underlying processing capacity across evolution from our common ancestor. Based on preliminary evidence for individual distinctiveness of drumming acoustic patterns in male chimpanzees, it has been proposed that these rhythmic cues may help the chimpanzees to recognize unseen conspecifics by their drumming behaviors (e.g. Arcadi et al., 1998). Although poorly studied and unclear, the question of the individual distinctiveness of rhythmic acoustic behaviors in animals remains essential to investigate this latter hypothesis.

In the present study conducted in a group of habituated wild chimpanzees living on the field site of Fongoli, South-East Senegal (Pruetz & Bertolani, 2007), we

investigated the individual rhythmic acoustic patterns generated by 4 females and 10 males when cracking a baobab fruit (Strydinos spp). In holding the fruit by its stem, this unimanual behavior consists of repetitively slapping the fruit on a hard surface (the ground, baobab tree's branch or root) in order to make its shell softer before opening and eating it. It resulted in bouts of long-distance rhythmic sounds which were recorded on site. Each bout of baobab fruit cracking was further analyzed to extract the individual inter-beat durations within the bout. Among the 151 bouts collected from 40 hours of video clips, we identified 3 types of rhythmic patterns produced by the chimpanzees: (1) Repetitive power beats in 8 subjets; (2) Repetitive sequences of 1 power beat + followed by either 1 rebound soft beat (in 9 subjects) or (3) + by 2 consecutive soft rebound beats (in 5 subjects). In addition, for each rhythmic patterns, the analyses of the individual averaged inter-beat durations across subjects showed significant differences between individuals.

These findings clearly indicate that cracking baobab fruit indirectly generated distinct inter-beat individual signatures in most of the chimpanzees. It is then not excluded that these rhythmic cues may help chimpanzees to recognize unseen conspecifics. Given the functional advantages of such potential long distance individual identifications, it constitutes an ideal candidate for selective pressure in favor of rhythmic processing prerequisites across primate evolution.

Acknowledgements

We are very grateful to William D. Hopkins, Dondo Kanté, Tama Sadiakhou, Wally & Syra Camara. Grants: Fondation Fyssen

- Arcadi, A. C., Robert, D. & Boesch, C. (1998). Buttress drumming by wild chimpanzees: temporal patterning, phrase integration into loud calls and preliminary evidence for individual distinctiveness. *Primates*, 39, 505–518.
- Cook, P., Rouse, A., Wilson, M. & Reichmuth, C. A. (2013). California sea lion (*Zalophus californianus*) can keep the beat: motor entrainment to rhythmic auditory stimuli in a non vocal mimic. *Journal of Comparative Psychology*, 127, 412–427.
- Fitch, W. T. (2006). The biology and evolution of music: A comparative perspective. *Cognition*, 100, 173–215.
- Pruetz, J. D., & Bertolani, P. (2007), Savanna Chimpanzees, Pan troglodytes verus, Hunt with Tools. Current Biology 17, 412–417.
- Ravignani, A. *et al.* (2016) 'What Pinnipeds Have to Say about Human Speech, Music, and the Evolution of Rhythm', *Frontiers in Neuroscience*, 10, 274.

TOPIC-OPENENDEDNESS: WHY RECURSION IS OVERRATED

IRIT MEIR*

*Corresponding Author: imeir@univ.haifa.ac.il Department of Communication Sciences and Disorders and Sign Language Research Lab, University of Haifa, Haifa, Israel

Humans can use language to refer to and describe endless varieties of situations, thoughts, ideas, and topics, including hypothetical situations and events that never happened. This capacity, referred to here as *topic-openendedness*, is a key feature distinguishing human languages from animal communication systems and any theory of language and language evolution should account for it. Recursion, the mechanism that provides language with the capacity for discrete infinity, can account for the fact that languages can create an infinite number of sentences from a finite set of words and rules. But it *cannot* account for the openendedness of the *contents* of those sentences. Therefore, the importance attributed to recursion as the *sole* mechanism that is *uniquely human* is overrated. We suggest that a key factor in explaining topic-openendedness is the nature of the linguistic symbols, the words, specifically their ability to extend their meanings beyond their basic meaning, to other, novel semantic domains, by means of cognitive processes such as metonymy and metaphor. Meaning extensions allow language users to apply a finite lexicon to an infinite number of situations and topics, and play a crucial role in explaining topic-openendedness.

1. What is special about human language: Topic-openendedness

In the past few decades, recursion has come to be regarded as one of the most fundamental properties of human languages and the human capacity for language. In Hauser et al. (2002) it was upgraded to the sole feature of FLN, the faculty of language in the narrow sense, which, according to the authors, consists of those features that are exclusively characteristic of human language, not shared by other human cognitive abilities or by other species. The reason for the importance attributed to recursion is that it is regarded as the property responsible for the openendedness of human language, its capacity for 'discrete infinity', which distinguishes human language from animal communication systems (ACSs).

However, discrete infinity, the ability to create an infinite number of sentences from a finite set of words and rules, is but one of the facets of language's

openendedness, and, we argue here, not the most important one. The importance of the openendedness of human language lies not only in its ability to create an infinite *number* of sentences, but also in what we can *convey* with these sentences. Language allows its users to refer to endless varieties of situations, thoughts, ideas, and topics. By using language, we can refer to and describe any topic that we feel a need to express, situations that are detached from the here-and-now (allowing for *displacement*, c.f. Hockett, 1960), hypothetical situations, situations that will not or cannot take place, and novel situations. *We can always use language to relate to new situations*. This ability, called here *topicopenendedness*,¹ stands in marked contrast to ACSs, that express information revolving around survival: food and feeding, predator-prey relations, mating and reproduction, and signaling social hierarchies (Hauser, 1996).

The paper makes several novel points. First, I argue that the uniqueness of human language lies in topic-openendedness as much as in discrete infinity. Recursion does not facilitate topic-openendedness, and therefore we should shift our focus of investigation and try to find factors that do contribute to it. Second, I suggest that a key factor in explaining topic-openendedness is the special nature of linguistic symbols, the words, especially their ability to take on meaning extensions by means of processes such as metonymy and metaphor. Though meaning extensions have been studied thoroughly in the semantic literature, their important role in facilitating topic-openendedness has not been previously acknowledged. I present three arguments to support these claims: (1) A thought experiment showing that recursion cannot account for topic-openendedness, but that meaning extensions can (section 2); (2) Evidence showing that recursion is not necessary to account for the openendedness of language, as languages can do without a syntactic mechanism for recursion and still express recursive thoughts (section 3); (3) Evidence from young languages showing that syntactic recursion is not found in early stages of a language, but even in its very early stages, a language can refer to novel and displaced situations, distinguishing it from ACSs (section 4). The concluding section (5) suggests a few factors that contribute to topic-openendedness, and should become the focus of future studies.

2. Topic-openendedness: the role of meaning extensions

All languages are characterized by topic-openendedness. Which factors contribute to this special ability? I propose that a central factor is the flexibility of the

¹ Topic-openendedness is similar, though not identical, to Katz's (1978) Principle of Effability, that states that anything that can be thought can be expressed in any human language. I thank Mark Aronoff for this point.

symbols that make up language, the words. Our ability to create and use symbols is at the heart of our special linguistic ability (Deacon, 1997, Bickerton, 2009). Words, vocal or gestural (manual) signals that are associated with (or represent) a concept (Pinker and Jackendoff, 2005), are crucial for using language to refer to entities of various kinds (humans, animals, objects), actions, states, emotions, abstract concepts and more. The ability to create new words, employed constantly and incessantly by language users, is the basis for our capacity to apply language to novel situations. Yet even a very large lexicon is finite at any given point in time and for any given individual, while the number of types of situations, entities and concepts we want to refer to is infinite. What happens when we encounter a situation which we have no words for? We can create new words, and we often do. But more often, we extend the meanings of the words we have to cover those new needs. So, interestingly, though the meaning of words is precise (a property which Hockett, 1960, refers to as 'semanticity' and lists as one of the design features of human languages), it is not altogether fixed. It is this flexibility that underlies our ability to refer to novel entities and situations.

I demonstrate the power of this flexibility through a thought experiment. Consider two languages, L_A and L_B . The two have almost identical vocabularies (very limited) and the same sentence structures. They differ in the following: only L_A has a recursive mechanism; it can embed a constituent within the same kind of constituent (in the example here, by means of a complementizer, but other means are possible too, of course), creating embedded structures of potentially infinite length. In addition, each word has precisely one meaning; neither polysemy nor any other meaning extensions is possible. L_B does not have a mechanism for recursion. However, it allows for polysemy, specifically metonymy and metaphor. Crucially, both languages are compositional and generative: both have the ability to combine words to create larger units, such as phrases and sentences.

nouns	verbs	adjectives	prepositions	COMP
man, woman, dog, head, house, sky, water, apple	eat, run, say, write walk	small, white, big	in, after, to	that (only in L _A)

Table 1. Vocabulary of both languages

Here are some situations that both languages can refer to: (1) *The man/woman/dog ate an apple.* (2) *The man/woman/dog run after the man/woman/dog– run in the house – run to the house/apple/man/woman/dog/water.* (3) *The woman/man/dog/house/apple is big/small/white.* (4) *The big/small/white woman/man/dog run/eat/read/write etc.* In addition, L_A can produce embedded structures such as: (5) *The woman/man said to the man/woman that ...*(any of the

above sentences). (6) *The woman said that the man said that the woman said that* ... (any of the above). So, formally, L_A is infinite: it can produce an infinite number of sentences. L_B , lacking the mechanism of recursion, cannot do that.

However, let's consider a situation where speakers of each of the languages want to refer to an animal other than a dog, say, a cat or a donkey. L_A does not provide its speakers with means of doing it, since it avoids polysemy or meaning shifts. L_B speakers, on the other hand, can use the word *dog* to refer to other animals as well, simply by extending its denotation, as small children often do. They can also create compounds such as small-dog for 'cat' and big-dog for 'donkey'. LA speakers can also create combinations like *small dog* and *big dog*, but they can only mean a small or a big dog respectively, since the word *dog* cannot be extended to other animals. L_B speakers can also create compounds like *sky-dog* and *water-dog* for 'a bird' and 'a fish' (or any other water animal) respectively. Similarly, the meaning of *apple* can be extended to other fruits and vegetables; the meaning of *head* can be extended to denote 'the top part of X', and then the spatial relation 'on'. So speakers of L_B can say *dog head house*, meaning 'the dog is on the house'. They can create new verbs such as *head-say* for 'think', watereat for 'drink', and new adjectives such as *head-big* 'smart', *head-small* 'stupid'. If they want to refer to something new, or clean, they can use the word *white* to refer to all three properties – white, new and clean. They can also use *white-sky* for 'clouds', white-man/woman for 'a good man/woman', white-dog for 'a sheep', and white-say to 'saying good things, praising'.

All these semantic shifts that L_B uses are very familiar. Languages use meaning extensions, such as metonymy and metaphor, to refer to new situations. There is nothing new in these examples. In fact, they are quite trivial; we are so used to them that they have escaped our radar when we try to explain what is so special about language. But the point is that these processes play a central role in our ability to use language to convey a novel concept or situation *-- any* concept or situation. Therefore, although L_A is infinite in the mathematical sense, it is rather limited when it comes to referring to new situations. L_B may be finite from a mathematical point of view, yet it is so much richer than L_A in terms of its expressive capabilities; it is not limited by its vocabulary to specific topics. The comparison between these two artificial languages shows that even with a very limited vocabulary, a communication system (L_B) can exhibit topic-openendedness, a communication system (L_A) is much less language-like, even if it mathematically infinite.

3. Expressing recursive thoughts: is syntactic recursion necessary?

Meaning extensions, and not recursion, can explain how we use a finite vocabulary to refer to an infinite number of situation and thoughts. Yet maybe

there are types of thoughts that can only be expressed by a recursive mechanism. For example, is recursion necessary for expressing complex, recursive thoughts, that is, propositions embedded in other propositions, another aspect of the infiniteness of language? It turns out that this is not the case. Recursive thoughts can be expressed by parataxis, putting one utterance after another, rather than by syntactic recursion or embedding. Evans and Levinson (2007, 443) show that content conveyed by an embedded structure (e.g. a conditional clause), can also be conveyed by non-embedded syntactic structures: "Consider that instead of saying, "If the dog barks, the postman may run away," we could say: "The dog might bark. The postman might run away." In the former case we have syntactic embedding. In the latter the same message is conveyed, but the "embedding" is in the discourse understanding - the semantics and the pragmatics, not the syntax." Similarly, the content of relative clauses can be expressed by parataxis. Instead of saying "The man who is watching TV is combing his hair", one can say "The man is watching TV, (the man) is combing his hair." (as old ISL signers do, see Dachkovsky 2017).

Similarly, in Al-Sayyid Bedouin Sign Language (ABSL), a village sign language of Israel that emerged in the Al-Sayyid community in the early 30s of the 20th century (Sandler et al. 2005, Meir et al. 2010), reported speech, another instance of recursive thoughts, is conveyed by several means, none of which involves syntactic recursion. One way is to cite the content of the speech, without explicitly indicating who said it. In a narrative told by a first generation signer of ABSL, he signs the following dialogue: - GUN GIVE-ME 'Give me the gun'; - SWEAR (BY-)GOD, NO 'I swear by God's name, no!' (Sandler, 2012). The signer does not embed the reported content in another clause. The fact that these utterances are instances of reported speech is indicated by his body posture and by the general context and shared background. There is no syntactic mechanism involved here, yet the semantic function of reported speech is conveyed.

Another way of expressing reported speech is by mentioning the speaker, and the content of the speech: FATHER: NO, STAY HOME 'Father [said]: no, you stay home'. Again, there is no syntactic embedding, as there is only one clause, and a noun preceding it. The embedding is clearly in the semantics or pragmatics. Everett (2012;196) mentions precisely the same mechanism for reporting speech in Wari', an Amazonian language.

A third generation ABSL signer uses the sign SAY to introduce reported speech. This mechanism is very similar to direct speech in English. FATHER SAY: WHY YOU LONG-TIME SEE NONE WHY? 'Your father said: "why haven't we seen you for such a long time, why?"

In all three cases, the signer expresses recursive thoughts, the embedding of one proposition (the content of the saying) in another (the saying event). But none of the ABSL mechanisms encodes syntactically the recursive nature of these

propositions. These data, and the data presented in Evans and Levinson (2007), provide evidence that syntactic recursion is not necessary for expressing recursive thoughts or messages (Jackendoff, 2011. See also Gil 2009 for Indonesian/Malay).

Furthermore, in many languages it is not clear that recursion leads to discrete infinity. In some languages, e.g., Kayardild (Evans & Levinson 2007, 442), recursion is limited to one cycle of application. Other languages use embedded structures very rarely, e.g. polysynthetic languages, in which the complexity resides in the morphology rather than in the syntax (see Evans, 2003; Mithun, 1984 for specific languages). Finally, even in languages where recursive structures are very common (e.g. English), center embedding beyond two levels is almost non-existent, and tail embedding usually does not exceed three or more levels in actual language use (Heine and Kuteva 2007, 297). Taking all the evidence provided in this section together, we conclude that recursion cannot carry the burden of accounting for the openendedness of human language.

4. The diachronic perspective: the view from young languages

We turn back to what we claim here is the central aspect of the openendedness of language, topic-openendedness. Further evidence for its centrality comes from novel and very young languages. Studies of such communication systems show that even at very early stages of their emergence, languages exhibit topic-openendedness. Even homesign systems, gestural communication systems invented by deaf children with no exposure to a conventional sign language, can do this (Goldin-Meadow 2003, 2005). The literature of pidgins and creoles is abundant with examples that show displacement (see e.g., Holm 1989). Studies of early stages of young sign languages in Israel show that 1st and 2nd generation signers tell stories of the history of their community, talk about diverse topics such as folk remedies, social security rights and health issues, and plan future actions and events (Sandler et al. 2005, Kastner et al. 2014, Meir et al. 2016. See also Ergin 2017 for Central Taurus Sign Language in Turkey).

These languages are often described as having very little syntax, and no mechanism for syntactic embedding. Other studies show that markers of embedding may develop over time. In some cases, embedding is marked by prosodic cues, e.g., a special facial expression and a forward body posture (see Sandler et al., 2011, for the development of systematicity in the prosodic marking in ABSL), or by means of shortening, both in time and in space, the movement of the predicate to mark it as secondary (dependent) to the main predicate in KQSL (Kastner et al., 2014). Similar findings are reported about Nicaraguan Sign Language (Senghas et al., 2016).

In other cases, overt markers of subordination can develop over time as well. Dachkovsky (2017) traced the development of relative clauses in three generations of ISL signers. The first generation did not have a consistent way of marking relative clauses. The second generation marked relative clauses by a specific facial expression – eye squint. In the third generation, a pointing sign, usually used as a demonstrative pronoun, developed into a marker of relative clause boundaries. Similar developments of embedded structures have been reported for many spoken languages (Heine & Kuteva, 2007, ch. 5).

The main point here is that syntactic marking of subordination often develops over time in languages. Languages often have a 'non-subordination' stage in their history, developing the means for overtly marking subordination over time, by means of grammaticalization (which involves phonetic, semantic and morpho-syntactic changes). Yet no language that I know of was reported as having a stage of 'topic-finiteness', that is, a stage in which the language users were confined to a closed set of topics. And it is this characteristic that makes it distinct from ACSs, that make it language. In other words, a language without recursion is still a language, whereas a language without topic-openendedness is not.

5. Topic-openendedness: symbols, meaning extensions and compositionality

So far I have established the claim that what makes a communication system a language is first and foremost topic-openendedness, and if we want to understand language we should try to account for that rather than for the ability for discrete infinity, which, as we saw, some languages may do without. From an evolutionary perspective, we have to explain how the capacity for topic-openendedness evolved.

I cannot make any specific suggestions to this mystery, and I know of no satisfactory suggested scenarios in the literature. What I would like to do here is to suggest some possible directions for exploration, by bringing to the forefront several properties that I believe enable topic-openendedness. If we can trace the evolution of any of these properties, we would make a great leap forward.

I suggest that two sets of properties are involved in this feat: the nature of the symbols and compositionality. The first has been the focus of this paper, the second will only be mentioned briefly here.

Humans have the capacity to create symbols, even without exposure to conventional symbols. This is evident in homesign systems, where deaf children invent signs that were not used by their caregivers (Goldin-Meadow, 2003), and by young sign languages, whose developing vocabulary is a testimony to this capacity (Meir et al., 2016). In the manual-visual modality, many of these signs are iconic; humans, unlike other primates, have the capacity to create iconic symbols (Sandler, 2010, 253). Humans may use iconic signs to refer to abstract notions by means of metonymy – depicting iconically an object, entity or action related to the abstract concept (e.g. wheat harvesting for 'year', and a handshake for 'holiday' in ABSL). In addition, humans extend the meanings of the words in

their lexicons to refer to and express novel situations, by means of metonymy and metaphor. For example, in a narrative of a first-generation ABSL signer, he uses the signs for SWORD, GUN, CUCUMBER metonymically, to refer to people: the man with the sword, the man with the gun, and the man that grows cucumbers. Metaphorical extensions are also evident in the creation of signs (see Taub, 2001, for an extensive analysis of metaphors in sign languages) and in creating meaning extensions, such as using the sign THIRSTY to mean 'want' in Algerian Jewish Sign Language (Meir et al., 2016).

Spoken languages are not different from sign languages in this respect. They also extend the use of their words to novel entities and situations. Any glance at any item in a dictionary would make this point. There are hardly any items that have only one sense. Even a grammatical word such as *the* has four senses or functions, with several sub-functions in each (Meriam-Webster online dictionary)²; the word *cat* has seven senses, and the word *run* has over 50 senses or meaning. The exact number of senses or meanings may be hard (or even impossible) to determine, but this is precisely the point: we use our finite set of words in novel and creative ways to refer to an infinite number of concepts and situations. This is a different facet of von Humboldt's famous saying that language "makes infinite use of finite means".

It is not clear how the ability for meaning extension developed in the course of language evolution. The study of ape gestures might be a possible starting point. It has been suggested that ape gestures show greater flexibility than facial expressions and vocalizations in terms of the behavioral contexts they occur in, and therefore the interpretation of gestures need to take into consideration a larger combinatorial context (Pollick and de Vall, 2007). Furthermore, ape gestures may show the buds of metonymy. When an ape scratches his body in a particular way to request to be scratched at this spot, the directed scratch is associated with the intended action by metonymy: it represents the entire event of scratching, including an agent different from the communicator (Pika and Mitani 2009; see also Hobaiter and Byrne 2014 for a large inventory of meaningful gestures in chimpanzees). These gestures are instructions for actions rather than referential symbols, but they nonetheless make use of metonymy. Though the path from these gestures to the wide use of meaning extension in human language is still a mystery, at least there is a possible evolutionary precursor to start from (Liebal and Call 2012).

The second set of properties that are needed to account for topic-openendedness are related to the notion of compositionality, our ability to combine symbols to create novel utterances whose meaning can be calculated on the basis of the meanings of their components and the way they are combined (Frege 1892). This ability is crucial in order to create a system with generative power. A good deal

² https://www.merriam-webster.com/dictionary/the

has been written about compositionality, and we will not develop this issue further here.

A specific kind of semantic relation which is based on compositionality is predication, the ability to 'say something about something', to attribute a property, an action or an event to a referent. Bickerton (e.g., 1990, 2009) has acknowledged and emphasized the importance of predication to human language and the evolution of syntax, and tried to suggest possible evolutionary precursors, but these are analogical (e.g. the behavior of foraging ants), and very different qualitatively from what we find in human languages. Others suggested that predication is a development within homo sapiens, built on neural machinery that underlies other systems, such as CV syllable structure (Carstairs-McCarthy 1999) or the inherent asymmetry between the hands when manipulating tools (Krifka 2008). While the origin of predication is crucial for achieving displacement, since it enables us to separate a protagonist from the event it is performing or undergoing, and to place a situation in time with respect to speech time and therefore to detach events form the here-and-now.

Related to predication is negation. As far as I know, among natural communication systems, only human languages have the ability to negate, to state that something does not exist or did not happen, an important aspect of topic-openendedness. Though ACSs may express related notions such as refusal or forbidding, these are still manipulative in nature and are not used as an assertion.

All the above properties – use of symbols, meaning extensions and compositionality, including predication and negation – are found at the very early stages of languages, enabling them to exhibit topic-openendedness, and setting them apart from other ACSs. If we want to characterize and understand human language and how it evolved, we should address these issues, rather than focus mainly on recursion.

- Bickerton, D. (1990). *Language and Species*. Chicago & London: The University of Chicago Press.
- Bickerton, D. (2009). Adam's Tongue: How Humans Made Language, How Language Made Humans. New York: Hill and Wang.
- Carstairs-McCarty, A. (1999). *The Origins of Complex Language*. Oxford: Oxford University Press.
- Dachkovsky, S. (2017). Grammaticalization of intonation in Israeli Sign Language: From information structure to relative clause relations. PhD dissertation, University of Haifa.
- Deacon, T. (1997). The Symbolic Species: The Coevolution of Language and the Brain. New York: Norton.
- Ergin, R. (2017). Central Taurus Sign Language: A Unique Vantage Point into Language Emergence. (Ph.D. Ph.D. Thesis), Tufts University. <u>https://search.proquest.com/openview/69f73091abbc97d1d2997408076c530</u> 1/1?pq-origsite=gscholarandcbl=18750anddiss=y
- Evans, N. (2003) Bininj Gun-wok: A Pan-dialectal Grammar of Mayali, Kunwinjku and Kune. Canberra: Pacific Linguistics.
- Everett, D. (2012). Language: the Cultural Tool. New York: Pantheon Books.
- Frege, G. (1892). On Concept and Object. Vierteljahrsschrift für wissenschaftliche Philosophie, 16: 192–205, Peter Geach (trans.), (1952) Translations from the Philosophical Writings of Gottlob Frege, Oxford: Blackwell, pp. 42–55.
- Gil, D. (2009). How much grammar does it take to sail a boat? In G. Sampson, D. Gil, and P. Trudgill (eds.), *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press, 19-33.
- Goldin-Meadow, S. (2003). The Resilience of Language: What Gesture Creation in Deaf Children Can Tell Us about How All Children Learn Language. New York: Psychology Press.
- Goldin-Meadow, S. (2005). What language creation in the manual modality tells us about the foundations of language. *Linguistic Review*, 2005, 22, 199-225.
- Hauser, M. (1996). *The Evolution of Communication*. Cambridge, MA: MIT Press.
- Hauser, M., Chomsky, N. and Fitch, T. (2002). The faculty of language: what is it, who has it, and how did it evolve. *Science* 298, 1569–1579.
- Heine, B. and Kuteva, T. (2007). *The Genesis of Grammar: A Reconstruction*. New York: Oxford University Press.
- Hobaiter, C. and Byrne, R. W. (2014). The meanings of Chimpanzee gestures. Current Biology 24, 1596–1600, <u>http://dx.doi.org/10.1016/j.cub.2014.05.066</u>
- Hockett, C. F. (1960). The origin of speech. Scientific American, 203, 88–111.
- Holm, J. (1989). *Pidgins and creoles*. Vol II. Camridge: Cambridge university Press.
- Jackendoff, R. (2011). What Is the Human Language Faculty? Two Views. Language, 87(3) 586-624.
- Kastner, I., Meir, I., Sandler, W. and Dachkovsky, S. (2014). The emergence of embedded structure: Insights from Kafr Qasem sign language. *Frontiers in Psychology*, 5, 1-15.
- Katz, J. (1978). Effability and translation. In F. Guenthmer, and M. Guenthner-Reutter (eds.), *Meaning and Translation: Philosophical and Linguistic Approaches*. New York: New York University Press, 191-234.
- Kocab, A., Senghas, A. and Snedeker, J. (2016). Recursion in Nicaraguan Sign Language. Proceedings of the annual meeting of the Cognitive Science Society. https://mindmodeling.org/cogsci2016/papers/0239/index.html

- Krifka, M. (2008). Functional similarities between bimanual coordination and topic/comment structure. In: Eckardt, Regine, Gerhard Jäger and Tonjes Veenstra, (eds), Variation, selection, development. Probing the evolutionary model of language change. Berlin: Mouton de Gruyter, 307-336.
- Liebal, K. and Call, J. (2012). The origins of non-human primates' manual gestures. *Philosophical Transactions of the Royal Society B*, 367(1585): 118– 128. doi: 10.1098/rstb.2011.0044
- Meir, I., Sandler, W., Padden, C. and Aronoff, M. (2010). Emerging Sign Languages. In: M. Marschark, and P. Spencer (Eds.) Oxford Handbook of Deaf Studies, Language, and Education, Vol. 2. Oxford: Oxford University Press, 267-280.
- Meir, I., Sandler, W., Ziv Ben-Zeev, Y., Lanesman, S. and Sarsour, M. (2016). How Does a Lexicon Emerge in a New Language? Insights from Village Sign Languages of Israel. *Iyunim be Safa ve-Hevra*, 8(1-2), 11-43. (in Hebrew).
- Mithun, M. (1984) How to avoid subordination. *Berkeley Linguistic Society* 10:493–509.
- Pika, S. and Mitani, J. C. 2009. The directed scratch: evidence for a referential gesture in chimpanzees? In Botha, Rudolf and Chris Knight (eds.) *The Prehistory of Language*. Oxford: Oxford University Press, 166-180.
- Pinker, S. and Jackendoff, R. (2005). The faculty of language: what's special about it? *Cognition* 95, 201–236.
- Pollick, A. S. and de Vall, F. B. M. (2007). Ape gesture and language evolution, *Proceedings of the National Academy of Sciences* 104:8184-8189.
- Sandler, W., Meir, I., Dachkovsky, S., Padden, C. and Aronoff, M. (2011). The emergence of complexity in prosody and syntax. *Lingua* 121, 2014-2033.
- Sandler, W., Meir, I., Padden, C. and Aronoff, M. (2005). The Emergence of grammar: Systematic structure in a new language. *Proceedings of the National Academy of Sciences* 102(7), 2661-2665.
- Sandler, Wendy. (2009). Symbiotic symbolization by hand and mouth in sign language. *Semiotica*, 174(1/4), 241-275.
- Sandler, Wendy. (2012). Dedicated gestures in the emergence of sign language. *Gesture*, 12(3), 265-307.
- Taub, S. (2001). Language from the Body. Iconicity and Metaphor in American Sign Language. Cambridge: Cambridge University Press.

DYNAMICAL PREFERENCES AND SELF-ACTUATED LANGUAGE CHANGES

JEROME MICHAUD*1,2

*Corresponding Author: jerome.michaud84@gmail.com
¹Department of Sociology, Uppsala University, Uppsala, Sweden
²Department of Physics and Astronomy, Uppsala University, Uppsala, Sweden

A puzzling fact about linguistic norms is that they are mainly stable, but the conventional variant sometimes changes. These transitions seem to be mostly S-shaped and, therefore, directed. Previous models have suggested possible mechanisms to explain these directed changes, mainly based on a bias favoring the innovative variant. What is still debated is what is the mechanism that leads to such a bias. In this paper we propose a refined taxonomy of mechanisms of language change and identify a family a mechanisms explaining self-actuated language changes. We exemplify this type of mechanism with the *preference-based* selection mechanism that relies on agents having dynamical preferences for different variants of the linguistic norm. The key point is that if these preferences can align through social interactions, then new changes can be actuated. We present results of a multi-agent model and demonstrate that the model produces trajectories that are typical of language change.

1. Introduction

An important question asked by Blythe and Croft (2012) is: 'How many qualitatively distinct possible mechanisms of language change are there?' (Blythe & Croft, 2012, p. 270). Based on the generalized theory of selection by Hull (1980, 2001, 2010) adapted to language evolution by Croft (2000) and by Baxter, Blythe, Croft, and McKane (2006), Blythe and Croft (2012) proposed a classification of mechanisms influencing language change. They distinguished four categories of mechanisms, namely neutral evolution (NE), neutral interactor selection (NIS), weighted interactor selection (WIS) and replicator selection (RS). The RS mechanism groups all mechanisms in which the variants are treated in a different manner by agents. Blythe and Croft (2012) argued that only RS can reliably account for Sshaped trajectories of change. In their model of RS, the innovative variant is given a selective advantage that causes its directed and S-shaped propagation through the population, but they did not provide an explanation for the origin of the shared advantage of the innovative variant and assumed it as given. It is unlikely that social, linguistic and cognitive factors, which can all induce a selective advantage for a variant, influence the dynamics of a language in the same way and lead, therefore, to qualitatively different selection mechanisms. The taxonomy proposed by

This paper is distributed under a Creative Commons CC-BY-ND license. DOI:10.12775/3991-1.074

306

Blythe and Croft (2012) has to be refined to account for this. With the help of a refined taxonomy, we investigate which type of selection mechanism can explain language change in absence of external (environmental) triggers. This is known as the *actuation problem* (Weinreich, Labov, & Herzog, 1968).

The model we propose is an extension of the utterance selection model (USM) by Baxter et al. (2006), where the state of an agent not only considers the probability distribution of use of a variant of a linguistic variable (Labov, 2001; Tagliamonte, 2012), but also takes into account a preference vector associated with the different variants. The idea to add preferences in the state of the agents has been used by Acerbi, Ghirlanda, and Enquist (2012) to explain fashion cycles. The influence of preferences on the behavior of agents is encoded in the updating rule of the agents rather than in the production rule as in Pierrehumbert, Stonedahl, and Daland (2014). The idea originates from Michaud (2014), where a learning rule based on game theory and related to reinforcement learning (Sutton & Barto, 1998; Roth & Erev, 1995) has been proposed. We have adapted this updating rule to the USM and proposed an alignment mechanism based on a bias for social adaptation based on probability matching (Gaissmaier & Schooler, 2008). This model is able to simulate different selection mechanisms and provides some insights on what properties are needed for self-actuated changes to occur. The results of our model are compared with the models of Stadler, Blythe, Smith, and Kirby (2016) and of Mitchener (2009).

2. Refinement of the RS mechanism

In Hull (1980, 2001, 2010) theory, a replicator (linguistic trait) is replicated by interactors (speakers) and selection can be active on both the replicator and the interactors. The replication process can be affected by random fluctuations (altered replication) and by selective advantage of some variant of the replicator (differential replication). The selection process is said to be replicator neutral is all the variants have the same chance to be selected.

The RS mechanism covers all the situations in which differential replication occurs, that is when a *fitness* advantage in the communication process is present. The origin or this selective advantage is not clearly specified by Blythe and Croft (2012), but they suggest that the differential weighting between variants should be construed as a social valuation of variants by speakers, see for example the socio-historical model of Labov (2001).

2.1. Heterogeneity of the RS mechanism

In Blythe and Croft (2012), RS is modeled by introducing a population-wide asymmetry between the different variants, favoring one over the other. This kind of *objective* advantage of one variant really questions the origin of such an asymmetry. How did all the speakers agreed that one variant is better than another? A simple explanation would be a kind of functional bias, which gives that variant an

advantage in communication. This is at odds with a social valuation explanation, since a social value should be agreed upon. In this section, we discuss the possible origins of selective advantage of variants of a replicator.

If the advantage of a variant is the same for all interactors one could call it *objective*. An objective advantage can originated in various biases, which can be cognitive, such as preferences for variants that are easier to process (Fedzechkina, 2014; Fedzechkina, Chu, Jaeger, & Trueswell, 2016), or they can be *linguistic*. Various linguistic biases such as: a regularization bias (Reali & Griffiths, 2009), a systematicity bias (Smith, Skarebela, & Tamariz, 2010), a naturalness bias (Finley & Badecker, 2007), an expressivity bias (Kirby, Tamariz, Cornish, & Smith, 2015), etc. have been proposed. These biases are objective in the sense that the linguistic system is the same for all speakers and their judgment of regularity, systematicity, naturalness, expressivity, etc. are similar. This type of biases are trivially aligned in the population. Contrarily to cognitive and linguistic factors, socio-cultural factors, such as social adaptation by probability matching, see, for example, (Gaissmaier & Schooler, 2008), are mainly subjective. The same variant can be socially valuated differently depending on the speech community it is used in. This means that such an effect should be attached to the speaker rather than to the community and a variant can propagate if there is an alignment of the individual valuations of this variant in the speech community. This type of *subjec*tive influences are of similar nature to the factors influencing trends and fashions formation and are qualitatively different from cognitive or linguistic biases. In the literature, notions such as variant prestige (Le Page & Tabouret-Keller, 1985; Cedergren, 1987; Tagliamonte, 2015) have been developed to account for this type of influences. It is the requirement for an alignment of the individual subjective valuations, which makes subjective and objective biases qualitatively different.

As we have seen, there is a qualitative difference between objective and subjective factors influencing the individual valuation of variants by the speakers. This implies that the RS mechanism should be refined.

2.2. Refined taxonomy

The fact that the taxonomy of Blythe and Croft (2012) has to be refined was also reached by Mühlenbernd and Michaud (2017) by weakening the assumption that the selective advantage was objective, i.e. shared by all the population. They find out that if the bias is subjective and randomly changing then there is no selection of a particular variant and the different variants coexist in the population.

The main distinction that has to be made between different RS mechanisms is whether the valuation of variants is the same for every agent (objective valuation) and depends on the social context and personal preferences of the interactors (subjective valuation). This distinction leads to two sub-categories of RS, an objective RS (ORS) and a subjective RS (SRS).

An additional useful distinction between RS mechanisms is related with the

time scale of change of the valuation. All types of biases evolve with time, but they do not evolve on the same time scale. On the one hand, cognitive biases evolve on a biological time scale and barely change on the time scale of language change. As a result, they can be well-approximated as static biases. On the other hand, social biases are highly variable and evolve on a faster time scale and cannot be considered as static but should dynamically evolve during language evolution. Linguistic biases are somewhat intermediary and can be considered either as static or as dynamic depending on which trait we are looking at. For lexical changes, structural biases stay roughly constant and can be considered as static. For grammatical changes the situation is different since a change in such a feature leads to a reorganization of the linguistic system itself, making these pressures dynamic.

With these two additional criteria, the RS mechanism proposed in Blythe and Croft (2012) should be refined into four different sub-mechanisms: (i) static and objective replicator selection (SORS); (ii) dynamic and objective replicator selection (DORS); (iii) static and subjective replicator selection (SSRS) and (iv) dynamic and subjective replicator selection (DSRS). It is important to note that a DSRS mechanism can be replicator neutral, whereas SORS, DORS and SSRS mechanisms are usually not.

2.3. Preference-based selection and self-actuation of changes

In this paper, we are looking for a replicator neutral mechanism of language change that can provide some insights into the actuation problem. For the changes to be self-actuated, we need a DSRS mechanism in which agents evaluate the different variants individually complemented by an updating rule for these valuations that leads to their population-wide alignment. We refer to individual valuation as *preferences*.

The preference-based selection mechanism proposed in this paper is a particular instance of DSRS. It assumes that the state of an agent is given by a probability distribution over the possible variants for a given trait and by a corresponding preference vector. This two-level description of the state of the agents is similar to the model by Acerbi et al. (2012) on the logic of fashion cycle, where preferences evolve and align when speakers interact. In order to understand how the mechanism works, let us assume that at a given point in time a variant of the replicator is consensual in the population. Due to altered replication, new variants can spontaneously emerge in the speech of individuals. If these newly introduced variants can influence the preferences of interactors (for example through social adaptation), then an interactor can grow a preference for a newly produced variant and starts to use it more. In other words, the interactor replicates her preferred variant differentially. Depending on the level of altered replication, the preference can either be reduced again or might be reinforced through random fluctuations. If it is reinforced, then other interactors will start to grow a preference for the same variant and a social valuation, i.e. a population-wide preference, starts to emerge.

As a result, the variant is differentially replicated and propagates through the population. Once the preferred variant has reached consensus, the preference stops to grow and starts to decrease due to altered replication. At some point, the preference for the variant is forgotten and a preference for a new variant can start to form. The functional form of the preference decay controls the period of stability of a given variant and acts as a constraint on when a new change can be actuated.

In this mechanism, the dynamics of the preferences shares a lot of properties with Stadler et al. (2016) momentum-based selection mechanism. For instance, both preferences and momentum align due to random fluctuations in the population (altered replication), the momentum is based on random time correlations in the usage of the variants, whereas preferences are based on random spatial correlations. Since the time correlation can be observed by all agents in the population, the momentum of a variant in a population tends to align, which leads to the differential replication of the variant. The same logic applies to preferences, since spatial correlations can also be measured by different agents. The effect of preferences and momentum on language use is similar in that it leads to differential replication of the favored variant. This differential replication increases both the temporal and spatial correlations of usage in the population, leading to the propagation of the variant. The main difference between our model and Stadler's model is that we do not need the concept of *age vector* or any structure in the population (Mitchener, 2009) to explain language change. In fact a weak bias for social adaptation by probability matching (Gaissmaier & Schooler, 2008) is sufficient to actuate new changes. This type of bias has been used by Jansson, Parkvall, and Strimling (2015) to model the evolution of creole languages. The overall dynamics of the preferences follows the same logic as that of fashion cycles (Acerbi et al., 2012). It starts with the joint emergence of a trait and its corresponding preference. The preference is then propagated through an prediction-driven instability (Mitchener, 2009) (because differential replication amplifies correlations of usage) and when its reaches a saturated state, the preference decreases until it becomes comparable with the preference for another variant. At this point a new change can be actuated.

3. Implementation of the preference-based selection mechanism

We implement the preference-based selection mechanism as an extension of the USM (Baxter et al., 2006) in which preferences are added. We call our model the USM with preferences (USMwP). In order to introduce preferences into the USM, we need (i) to modify the state space of the agents to include preferences; (ii) to define the influence of preferences on the behavior of agents and (iii) to define the preferences dynamics. For a detailed description of the model, see Michaud (2017) (USM without preferences) and the supplementary material. The first step is the easier one, since we only need to associate to each agent a preference vector π of length V (the number of variants) and whose components take values be-

tween 0 and 1. This complements the frequency vector x from which utterances u are produced. For the preferences domain, we could have chosen value between -1 and 1 as in Acerbi et al. (2012), but such a definition would only complicates the resulting accommodation rule, since we base our work on the learning rule used in Michaud (2014).

In Michaud (2014), the choice of a language is associated with the mixed strategy to play an abstract coordination game (Cooper, 1998) and the learning rule is based on lateral inhibition, that is, successful variants are reinforced and unsuccessful variants are penalized. This type of strategy has been used in naming games experiments (Beuls, Steels, & Höfer, 2012; Steels & Loetzsch, 2012; Steels, 2007). In Michaud (2014), the penalization parameter is formed by the product of a *objective cost* and a *subjective bias*. In our model, we define a *penalization vector* $\alpha := 1 - \pi$ from the preference vector, i.e. the variants which are preferred more are penalized less during the learning process.

The learning rule takes the form

$$x_{v}^{(i),n+1} = x_{v}^{(i),n} + u_{v}^{(j),n} \sum_{w=1}^{V} \alpha_{w}^{(i),n} x_{w}^{(i),n} - \alpha_{v}^{(i),n} x_{v}^{(i),n},$$
(1)

where the index n refers to the discrete time, the index i refers to the identity of the agent and the subscript v, w refers to the specific variants.

The last missing component of the implementation is the preferences' dynamics. We consider an updating rule for the preference vector π based on a bias for social adaptation (Gaissmaier & Schooler, 2008). To implement such a bias locally, agents have to be aware of the averaged speech of their neighbors. Let V_i be the set of neighbors of *i* on the network underlying the USM and let $U^{(i)}$ be the average utterance of the neighbors of an agent *i*. We define the updating rule for the preference vector $\pi^{(i)}$ by

$$\boldsymbol{\pi}^{(i),n+1} = \boldsymbol{\pi}^{(i),n} - \boldsymbol{\mu}(\boldsymbol{x}^{(i),n} - \boldsymbol{U}^{(i),n}), \tag{2}$$

where μ is a positive parameter controlling the rate of change of the preferences, i.e. the strength of differential replication. The preference change in the direction of the difference $U^{(i),n} - x^{(i),n}$, aiming to align the behavior of an agent with the average behavior of her neighborhood. Therefore, Equation (2) implements of a social adaptation bias (Gaissmaier & Schooler, 2008). The choice to compare the internal representation $x^{(i)}$ with averaged utterance U^i is motivated by the fact that we want to update our knowledge based on perceived information and since we do not have access to the information contained inside the brain of other agents, we must rely on their production. The parameter μ controls the strength of the preferences influence on the dynamics. According to the updating rule (2), the preferences of underused variants are increased and preferences of the overused variants are decreased. It is important to note that if one variant is conventional, then the corresponding preference can only decrease, since it is not possible to overuse a categorical variant. Due to altered replication, the preference for innovative or erroneous variants can stochastically build up. For instance, the updating rule (2) also accounts for a *surprise* effect. Since the preference for a conventional variant can only decrease, at some point its preference will become comparable to that of another variant.

4. Results and discussion

We test our model on a regular network of degree k = 3 mad of N = 20 agents. In order for no interactor selection to be active, we set the attention parameter to h = 0.5 for all agents and set $\lambda = 0.1$, for a detailed discussion of the influence of these parameter see Michaud (2017) and supplementary material.



Figure 1. Results for the application of time series measures. For each measures, 10 logarithmically spaced values for q and μ have been selected and, for each parameter combination, 100 runs of length 10000 network updates have been performed. The dashed vertical line corresponds to the critical q_* value of the USM with constant preferences $\alpha = 0.5$ for all agents and for all variants.

In these settings, we vary the strength of altered replication (controlled by q, see Equation (S1) in supplementary material) and the strength of differential replication (controlled by μ). For each combination of the parameters, we compute six measures of the time series generated, see Kauhanen (2016) and supplementary material. Results are displayed in Fig. 1.

When altered replication is increased dominance, monotonicity and smoothness decrease, while the number of shifts first increases and then decreases and while the length of shift increases. If altered replication is low, only one variant is used and altered replication is unlikely to generate a change. For larger values of q, altered replication is sufficient to trigger changes, but this changes tends to be highly stochastic (low logisticness) and slow (long length of shifts) in the absence of differential replication. When altered replication becomes too large, all the variants coexists and shifts become less likely.

The main effect of preferences, i.e. of differential replication, is to stabilize existing conventions and to speed up changes when they occur. In Fig. 1, this implies that the dominance, monotonicity, smoothness measures increase with μ . Shorter shifts also tends to be more logistic and the logisticness measure increases with differential replication. We also observe that the shifts become more frequent as well as shorter. In the high altered replication regime, high preferences leads to frequent and fast changes.

Our results demonstrate that the USMwP can predict various behavior of language change. Fig. 1 can be decomposed into four quadrants. The lower left quadrant, where altered and differential replication are low, accounts for stable features of language that are unlikely to change. The lower right quadrant, where altered replication is high but differential replication is low, accounts for highly variable features that coexist in the linguistic system. The upper left quadrant, where altered replication is low but differential replication is high, accounts for features that are relatively stable but can occasionally change in a directed and S-shaped manner. Finally, the upper right quadrant, where altered and differential replications are high account for features which change quickly and in a directed manner. This quadrant accounts for fashion like features that can change rapidly, similarly to the results of Acerbi et al. (2012).

To conclude, we have demonstrated that the DSRS mechanism proposed in this paper is able to simulate typical language change and what controls the shape and frequency of transitions is the competition between altered and differential replication.

Acknowledgements

The author has been partially funded by the Swiss National Science Foundation, grant number P2GEP2_159156. The author would like to thanks Richard A. Blythe, Roland Mühlenbernd and Anna Jon-And for interesting discussions.

References

- Acerbi, A., Ghirlanda, S., & Enquist, M. (2012). The logic of fashion cycles. *PloS one*, 7(3), e32541.
- Baxter, G. J., Blythe, R. A., Croft, W., & McKane, A. J. (2006). Utterance selection model of language change. *Physical Review E*, 73(4), 046118.
- Beuls, K., Steels, L., & Höfer, S. (2012). The emergence of internal agreement systems. *Experiments in Cultural Language Evolution. John Benjamins, Amsterdam.*
- Blythe, R. A., & Croft, W. (2012). S-curves and the mechanisms of propagation in language change. *Language*, *88*(2), 269–304.
- Cedergren, H. (1987). The spread of language change: Verifying inferences of linguistic diffusion. *Language spread and language policy: Issues, implications, and case studies*, 45–60.
- Cooper, R. (1998). Coordination games. Cambridge University Press.
- Croft, W. (2000). *Explaining language change: An evolutionary approach*. Pearson Education.
- Fedzechkina, M. (2014). *Communicative efficiency, language learning, and language universals.* Unpublished doctoral dissertation, University of Rochester.
- Fedzechkina, M., Chu, B., Jaeger, T. F., & Trueswell, J. (2016). Processing preferences shape language change. In S. Roberts, C. Cuskley, L. McCrohon, L. Barceló-Coblijn, O. Fehér, & T. Verhoef (Eds.), *The evolution of language: Proceedings of the 11th international conference (evolangx11).*
- Finley, S., & Badecker, W. (2007). Towards a substantively biased theory of learning. In *Berkeley linguistics society* (Vol. 33, pp. 142–154).
- Gaissmaier, W., & Schooler, L. J. (2008). The smart potential behind probability matching. *Cognition*, *109*(3), 416–422.
- Hull, D. L. (1980). Individuality and selection. Annual review of ecology and systematics, 11(1), 311–332.
- Hull, D. L. (2001). Science and selection. Cambridge Un. Press, Cambridge.
- Hull, D. L. (2010). Science as a process: an evolutionary account of the social and conceptual development of science. University of Chicago Press.
- Jansson, F., Parkvall, M., & Strimling, P. (2015). Modeling the evolution of creoles. *Language Dynamics and Change*, 5(1), 1–51.
- Kauhanen, H. (2016). Neutral change. Journal of Linguistics, 1-32.
- Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, 141, 87–102.
- Labov, W. (2001). Principles of linguistic change volume 2: Social factors (Vol. 29). Blackwell.
- Le Page, R. B., & Tabouret-Keller, A. (1985). Acts of identity: Creole-based

approaches to language and ethnicity. CUP Archive.

- Michaud, J. (2014). On the emergence of bilingualism in a communication "ALL" task as a result of competition between social conformism and language simplification. In E. A. Cartmill, S. Roberts, H. Lyn, & H. Cornish (Eds.), *The evolution of language: Proceedings of the 10th international conference (evolangx)* (pp. 185–192). World Scientific.
- Michaud, J. (2017). Continuous time limits of the utterance selection model. *Phys. Rev. E*, 95, 022308.
- Mitchener, W. G. (2009). A stochastic model of language change through social structure and prediction-driven instability. *language acquisition*, 24, 25.
- Mühlenbernd, R., & Michaud, J. (2017). The utterance selection model and different types of replicator selection. In 23 rd international conference on historical linguistics.
- Pierrehumbert, J. B., Stonedahl, F., & Daland, R. (2014). A model of grassroots changes in linguistic systems. arXiv preprint arXiv:1408.1985.
- Reali, F., & Griffiths, T. L. (2009). The evolution of frequency distributions: Relating regularization to inductive biases through iterated learning. *Cognition*, 111(3), 317–328.
- Roth, A. E., & Erev, I. (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and economic behavior*, 8(1), 164–212.
- Smith, A., Skarebela, B., & Tamariz, M. (2010). Exploring the nature of a systematicity bias: an experimental study. In A. Smith, M. Schouwstra, B. de Boer, & K. Smith (Eds.), *The evolution of language: Proceedings of the 8th International Conference (EVOLANG8)* (pp. 289–296). World Scientific.
- Stadler, K., Blythe, R. A., Smith, K., & Kirby, S. (2016). Momentum in language change. *Language Dynamics and Change*, 6(2), 171–198.
- Steels, L. (2007). The recruitment theory of language origins. In *Emergence of communication and language* (pp. 129–150). Springer.
- Steels, L., & Loetzsch, M. (2012). The grounded naming game. *Experiments in cultural language evolution*, *3*, 41–59.
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction* (Vol. 1). MIT press Cambridge.
- Tagliamonte, S. A. (2012). Variationist sociolinguistics: Change, observation, interpretation (Vol. 40). John Wiley & Sons.
- Tagliamonte, S. A. (2015). *Making waves: The story of variationist sociolinguistics.* John Wiley & Sons.
- Weinreich, U., Labov, W., & Herzog, M. I. (1968). Empirical foundations for a theory of language change. In W. P. Lehmann & Y. Malkiel (Eds.), *Directions for historical linguistics* (pp. 95–195). University of Texas Press.

THE PREVALENCE OF REPAIR IN STUDIES OF LANGUAGE EVOLUTION

ASHLEY MICKLOS^{*1,2}, VINICIUS MACUCH SILVA³, and NICOLAS FAY⁴

*Corresponding Author: ashleymmicklos@gmail.com
¹Linguistics Department, University of California, San Diego, USA
²Language and Cognition Department, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
³Department of Linguistics, University of Tübingen, Germany
⁴School of Psychology, University of Western Australia, Perth, Australia

While studies of language evolution have themselves evolved to include interaction as a feature of interest (Healey et al, 2007; Tamariz et al, 2017; Fay et al, 2017; Byun et al, in press), many still fail to consider just what interaction offers emerging communication systems. That is, while it's been acknowledged that face-to-face interaction in communication games is beneficial in its approximation of natural language use (Macuch Silva & Roberts, 2016; Nölle et al, 2017), there remains a lack of detailed analysis of what this type of interaction affords participants, and how those affordances impact the evolving language. To this end, here we will expose one particular process that occurs in interaction: repair, or the processes by which we can indicate misunderstanding and resolve problems in communication (Schegloff, Jefferson, & Sacks, 1977; Jefferson, 1972). Though it is often not explicitly analyzed, repair is a relevant aspect of interaction to consider for its effects on the evolution of a communication system as well as how it demonstrates the moment-to-moment processing and negotiation of alignment in emerging communication.

We present data from various studies of language evolution in which we document how repair is carried out, the types of repair present, and their effect on novel signaling. All studies in this collection utilized referential communication tasks – some iterated over simulated generations and other repeating interactions between two individuals. However, they differ in modality (of stimuli and communication). The data collection includes: silent gesture communication of written nouns and verbs; non-linguistic vocalizations and

gestures communication of either images or sounds; graphical communication of written nouns; and, text-based communication of tangram stimuli. Crucially, the stimuli are meant to be easily confusable. Repairs were categorized according to Dingemanse and Enfield's (2015) schema on other-initiated repair (OIR), or when a Matcher (Receiver) initiates a repair sequence that the Director (Sender) is meant to resolve.

In experiments with face-to-face interaction, repair sequences account for at least 20% of all turns. Even when participants are not face-to-face yet the task affords contingent, bidirectional communication, repair sequences take up almost 10% of all turns. These instances of repair are not trivial. We provide a descriptive analysis of how repairs are performed, and the outcomes of these repair sequences, which promote efficient, informative signals. The prevalence of these interactional sequences across studies of language evolution, and indeed across modalities, demonstrates that users of a novel communication mode make use of repair for the purposes of overcoming miscommunication and to establish alignment (with one another and to signal-meaning matches that eventually conventionalize).

Moreover, it has been argued that feedback can boost communicative efficiency, but not communication success (Fay et al, 2017). Results from a meta-analysis show that repair does indeed improve communicative efficiency when compared with non-repair trials or conditions, resulting in shorter, less complex gestural signals, more abstract and less complex drawings, and more concise textual descriptions. However, the effect of repair on communicative success has been less clear. We suggest that the affordances of certain repair types are more likely to improve communication success over other types (e.g. repairs that improve on prior signals, rather than lead to new innovations), and that repair sequences have more downstream effects on accuracy as opposed to *in-situ* success.

A goal of this review is to draw attention to the prevalence of repair in experimental investigations of language evolution. Repair, as a pragmatic universal (Evans and Levinson, 2009; Dingemanse et al, 2015) underlying human communication, is a mechanism that promotes efficient and successful communication. We find repair is ubiquitous across modalities and - even when not being directly tested - it is a factor that arises in, and affects the processes of, emergent communication. More broadly, we hope to call attention to not only the need to consider interaction as an ecologically valid site for language evolution and use, but also to consider the specific mechanisms within interaction that drive language to be structured as it is.

Acknowledgements

AM thanks her co-authors, and members of the LEvInSoN group at the Max Planck Institute for Psycholinguistics for their feedback. VMS thanks his Master's thesis committee: Seán G. Roberts, Judith Holler, and Asli Özyürek.

References

- Byun, K.S., de Vos, C., Bradford, A., Zeshan, U., & Levinson, S.C. (In Press). First encounters: Repair sequences in cross-signing. *Topics in Cognitive Science*.
- Dingemanse, M., & Enfield, N.J. (2015). Other-initiated repair across languages: Towards a typology of conversational structures. *Open Linguistics* 1 (1). 96–118.
- Dingemanse, M., Roberts, S. G., Baranova, J., Blythe, J., Drew, P., Floyd, S., ... Enfield, N. J. (2015). Universal principles in the repair of communication problems. *PLoS ONE*, 10(9), 1–15.
- Evans, N., & Levinson, S. C. (2009). The myth of language universals: language diversity and its importance for cognitive science. *The Behavioral and Brain Sciences*, 32(5), 429-448-494.
- Fay, N., Walker, B., & Swoboda, N. (2017). Deconstructing social interaction: The complimentary roles of behavior alignment and partner feedback to the creation of shared symbols. *Proceedings of the 39th Annual Meeting of the Cognitive Science Society*, July 26-29, London, United Kingdom.
- Healey, P.G.T., Swoboda, N., Umata, I., & King, J. (2007). Graphical language games: interactional constraints on representational form. *Cognitive Science*, 31 (2), 285–309.
- Jefferson, G. (1972). Side sequences. In D. Sudnow (Ed.), *Studies in Social Interaction* (pp. 294-338). The Free Press: New York.
- Macuch Silva V. & Roberts S. (2016) Exploring the Role of Interaction in the Emergence of Linguistic Structure. In S. Roberts & G. Mills (Eds.) Proceedings of EvoLang XI, Language Adapts to Interaction Workshop, 21 March, 2016.
- Nölle, J., Staib, M., Fusaroli, R., & Tylén, K. (2017). Informational bottlenecks lead to co-evolution of categories and systematicity in an emergent communication system. *Proceedings of Protolang 5*, Barcelona, September 2017.
- Schegloff, E.A., Jefferson, G., & Sacks, H. (1977). The preference for selfcorrection in the organization of repair in conversation. *Language*, 53 (2), 361–382.
- Tamariz, M., Roberts, S. G., Mart, J. I., & Santiago, J. (2017). The Interactive Origin of Iconicity. *Cognitive Science*, 1–16.

THE EMERGENCE OF ADJACENCY PAIRS: NO EVIDENCE IS BETTER THAN NEGATIVE EVIDENCE

GREGORY MILLS

g.j.mills@rug.nl

Centre for Language and Cognition, University of Groningen, Netherlands

When participants use dialogue in joint activities, they rapidly converge on idiosyncratic referring conventions. Convergence is inherently interactive, relying on participants providing each other with both positive and negative evidence of understanding (Clark, 1996; Healey, 2007).

In addition to securing reference, interlocutors also need to coordinate on the timing and sequencing of their contributions. Dialogue is replete with procedural expressions that establish who performs which action, when the action should be performed, and how initiation and completion of the action should be signaled, e.g. "when I've done x, do y"; "wait a moment before doing y"; "let's start again, but this time you do y"; "hold on I need to tell you x first"; "do x and then tell me when you're done"; "you do x and then I do y".

Recent work has demonstrated that participants rapidly establish new temporal expressions (Verhoef et al, 2016). Moreover, when temporal expressions are used interactively by participants to coordinate their actions with those of their partner, these expressions rapidly become conventionalized within novel adjacency pairs (Schegloff, 2007; Mills, 2014; Fusaroli et al., 2014). However, it is currently unclear how the central dialogue mechanisms of positive and negative feedback and alignment contribute towards conventionalization.

To investigate how procedural coordination develops, we report a computermediated "alien language" task which prevents participants from using natural language by restricting typing to a limited set of keys. However, in contrast to canonical alien language tasks which elicit referring expressions for describing referents, this task elicits procedural expressions for performing sequences of actions.

For example, on a typical trial, the task of the Director might be to instruct the Matcher that they have to perform the following 5-step sequence:

- (1) Matcher has to press A;
- (2) Matcher has to press S
- (3) Director and Matcher have to press F simultaneously
- (4) Director has to press J at the same time as Matcher presses **D**
- (5) Matcher has to press D

On each trial, the target sequence (i.e. **A**, **S**, **F**, **J**, **D**), valence (whether both perform the action together or only one participant performs the action), and congruence (whether both perform the same or different action) are determined randomly by the server. This presents participants with the recurring procedural coordination problem of communicating and then successfully performing a wide variety of sequences of actions, without using natural language. In order to test the putative role of positive and negative feedback on how coordination develops, dyads were assigned to one of 4 conditions:

- (1) Positive feedback: participants could send Y for yes
- (2) Negative feedback: participants could send N for no
- (3) Positive and Negative feedback: participants could send both Y and N
- (4) No feedback: participants were blocked from sending Y or N

Participants who could provide both positive and negative feedback correctly solved more trials in shorter time, confirming the basic predictions of the grounding model (Clark, 1996). Surprisingly, participants who could signal negative evidence of understanding with "N" performed worse than participants who were blocked from providing any feedback. We argue this pattern is due to the intrinsic ambiguity of "no": it signals but does not diagnose the problem, and gives no indication what the next relevant action should be. By contrast, the doubly-blocked participants are forced ab-initio to develop new, and consequently more robust, routines for establishing and sustaining procedural coordination.

All four conditions showed equal levels of alignment, but closer inspection showed that, overall, participants aligned *more* in unsuccessful trials, contradicting the interactive alignment model (Pickering and Garrod, 2009). We argue this is due to participants using alignment as a repair strategy: if participants know the next action, they perform that next action (Clark 1996). However, when participants do not understand the contributions of their partner, one of the default strategies is to repeat the actions of the interlocutor in order to establish a basic level of coordination that can then serve as the scaffolding for coordinating on more complex and complementary adjacency pairs.

References

Clark, H. H. (1996). Using language. Cambridge university press.

- Fusaroli, R., Rączaszek-Leonardi, J., & Tylén, K. (2014). Dialog as interpersonal synergy. New Ideas in Psychology, 32, 147-157.
- Garrod, S., & Pickering, M. J. (2009). Joint action, interactive alignment, and dialog. Topics in Cognitive Science, 1(2), 292-304.
- Healey, P. G., Swoboda, N., Umata, I., & King, J. (2007). Graphical language games: Interactional constraints on representational form. Cognitive Science, 31(2), 285-309.
- Mills, G. (2011, January). The emergence of procedural conventions in dialogue. In Proceedings of the Cognitive Science Society (Vol. 33, No. 33).
- Mills, G. J. (2014). Dialogue in joint activity: complementarity, convergence and conventionalization. New ideas in psychology, 32, 158-173.
- Schegloff, E. A. (2007). Sequence organization in interaction: Volume 1: A primer in conversation analysis (Vol. 1). Cambridge University Press.
- Verhoef, T., Walker, E. & Marghetis, T. (2016) Cognitive biases and social coordination in the emergence of temporal language. Proceedings of the 38th Annual Conference of the Cognitive Science Society. (pp. 2615-2620) Austin,TX: Cognitive Science Society.

DIFFERENTIAL RATES OF CHANGE IN CONSONANT AND VOWEL SYSTEMS

STEVEN MORAN^{*1} and ANNEMARIE VERKERK²

*Corresponding Author: steven.moran@uzh.ch ¹University of Zurich ²Max Planck Institute for the Science of Human History

Consonants and vowels are processed differently and they seem to have distinct neural representations (Caramazza et al. 2000). Böe et al. (2017) insist that vowel-like systems must be inferred to the last common ancestor of Baboons and humans, 25mya. Unlike vowels, however, consonants appear to be a later innovation in the communication systems of Hominids. Primates, including chimpanzees and orangutans, employ a repertoire of voiceless calls (so-called raspberries), which show homology with voiceless consonants (Lameira et al. 2014). During the course of human evolution, smaller orofacial cavities, increased neuro-cognitive abilities, and more precise motor control of the articulators led to greater phonetic variation, particularly among consonants, which have become phonologized in many ways in different language families. In comparison to vowels, there are over three times as many consonant phonemes in the world's languages. Their number and diversity ranges greatly, from 6 in Rotokas to over 90 in !Xu (Maddieson 1984); compare vowel systems which range in size from 2 to 14. Why are there are so many more consonants in the world's languages?

The answer to this question is complex, with factors involving a need for increased number of lexical contrasts in order to accommodate a growing vocabulary throughout the evolution of language, and the greater possibility for consonants rather than vowels to increase the number of contrastive sounds in a language through secondary articulations. Two strands of evidence support this conclusion. First, comparing a database of proto-language reconstructions (Marsico et al, accepted; n=100) with modern languages in UPSID (Maddieson 1984), Marsico (1999) notes an increase in the number of consonants in modern

phonological inventories (even though there are no great differences between reconstructed and synchronic inventories). Second, in a large sample of phonological inventories described with a rich distinctive feature set (Moran, 2012; n=1672), consonants are more often reported as marginal or borrowed than vowels (Moran et al, 2014). Both observations are in line with the need to increase vocabulary through the expansion of speech sound inventories via both vertical and horizontal transmission of languages.

Here we test whether six languages families (Arawakan (language sample n=38), Austronesian (83), Bantu (114), Indo-European (58), Pama-Nyungan (134), and Tupi-Guarani (30); references to phylogenetic tree sets below) show larger rates of change for consonant inventory size as compared to vowel inventory size using phylogenetic comparative methods. Our results suggest that within the reconstructable timescale of language families in our sample, rates of change are in fact larger for consonant inventory sizes are generally closer to the mean of the range rather than expert reconstructions of proto-languages, which warrants a closer evaluation using directional models of feature change. We also compare the fit of evolutionary models for continuous and discrete features to see which best accounts for phonological change.



Figure 1: Box plots of the ranges of vowel and consonant inventory size in the language samples used for phylogenetic ancestral state estimations. P gives proto-language reconstruction from Marsico (1999). R gives ancestral state estimation. * indicates whether the rate of change of vowels or consonants is faster.

References

- Böe L.-J., Berthommier F., Legou T., Captier G., Kemp C., Sawallis T.R., Becker, Y., Rey, A., & Fagot, J. (2017). Evidence of a Vocalic Proto-System in the Baboon (Papio papio) Suggests Pre-Hominin Speech Precursors. *PLoS ONE* 12(1): e0169321.
- Caramazza, A., Chialant, D., Capasso, R., & Miceli, G. (2000). Separable processing of consonants and vowels. *Nature*, 403(6768), 428.
- Lameira, A. R., Maddieson, I., & Zuberbühler, K. (2014). Primate feedstock for the evolution of consonants. *Trends in Cognitive Sciences*, 18(2), 60-62.
- Maddieson, I. (1984). Patterns of Sounds. Cambridge University Press.
- Marsico, E. (1999). What can a database of proto-languages tell us about the last 10,000 years of sound changes? In *Proceedings of the XIVth International Congress of Phonetic Sciences* (pp. 1-7).
- Marsico, E., Flavier, S., Verkerk, A., & Moran, S. (Accepted). BDPROTO: A Database of Phonological Inventories from Ancient and Reconstructed Languages. To appear in Proceedings of the 11th Language Resources and Evaluation Conference (LREC).
- Moran, S. (2012). Phonetics Information Base and Lexicon. PhD thesis. University of Washington.
- Moran, S., McCloy, D., & Wright, R. (Eds.) (2014). PHOIBLE Online. Leipzig: Max Planck Institute for Evolutionary Anthropology. See: https://github.com/phoible/dev

References for phylogenetic tree sets mentioned in abstract:

- Bouckaert, R., Lemey, P., Dunn, M., Greenhill, S. J., Alekseyenko, A. V., Drummond, A. J., Gray, R. D., Suchard, M. A. & Atkinson, Q. D. (2012). Mapping the origins and expansion of the Indo-European language family. *Science*, 337, 957-960.
- Bowern, C. & Atkinson, Q. D. (2012). Computational phylogenetics and the internal structure of Pama-Nyungan. *Language*, *88*(4), 817-845.
- Gray, R. D., Drummond, A. J. & Greenhill, S. J. (2009). Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science*, 323, 479-483.
- Grollemund, R., Branford, S., Bostoen, K., Meade, A., Venditti, C. & Pagel, M. (2015). Bantu expansion shows that habitat alters the route and pace of human dispersals. *Proceedings of the National Academy of Sciences*, *112*(43), 13296-13301.

- Michael, L., Chousou-Polydouri, N., Bartolomei, K., Donnelly, E., Wauters, V., Meira, S. & O'Hagan, Z. (2015). A Bayesian phylogenetic classification of Tupí-Guaraní. *LIAMES*, 15(2), 193-221.
- Walker, R. S. & Ribeiro, L. A. (2011). Bayesian phylogeography of the Arawak expansion in lowland South America. *Proceedings of the Royal Society B*, 278, 2562-2567.

THE CULTURAL EVOLUTION OF SPATIAL MODULATIONS IN ARTIFICIAL SIGN LANGUAGES

YASAMIN MOTAMEDI*¹, MARIEKE SCHOUWSTRA², JENNIFER CULBERTSON², KENNY SMITH², SIMON KIRBY² *Corresponding Author: y.motamedi@ucl.ac.uk ¹Experimental Psychology, University College London, London, UK ²Centre for Language Evolution, University of Edinburgh, Edinburgh, UK

Sign languages use a range of linguistic tools to denote the relationship between a predicate and its arguments, some which are shared with spoken languages (e.g. lexical contrasts and word order). However, many sign languages also make use of modality-specific spatial modulations to denote who does what to whom. The most common manifestation cross-linguistically of such spatial modulations is referential indexing, where animate arguments are indexed with a location in space and those indexed locations are referred to over a stretch of discourse to refer to the same argument. A characteristic example of this is given in 1); the man and the woman are represented by indexed locations—a and b, respectively—and the verb ask moves between the referenced locations.

1. MAN a WOMAN b a ASK bThe man asks the woman'

Though spatial modulation of this kind is widespread cross-linguistically, and has been considered near-universal in sign languages, recent study of emerging sign languages suggests that systematic spatial reference does not emerge fully formed but emerges and systematizes over generations of a new language (Padden et al., 2010; Kocab et al., 2014). We set out to examine how a tool that relies heavily on the iconic use of space systematizes over time, and whether an understanding of this process can shed light on the debate over the linguistic nature of spatial modulations (Cormier et al., 2015; Liddell, 2003; Lillo-Martin and Meier, 1999).

We present an experimental investigation into the emergence and evolution of spatial modulations in the manual modality, asking participants to communicate

about a set of events involving one or more animate arguments. The experiment took the form of a silent gesture task, in which hearing participants with no knowledge of sign language had to communicate a set of events using only gesture. The task combined both interaction, in the form of a director-matcher task, and iteration, with pairs of participants organized in transmission chains. A pair of participants was first trained on a set of gestures produced by a participant in the previous generation, and then in testing had to communicate about the same events, taking turns to produce (direct) and interpret (match) a gesture. The gestures they produced were then transmitted as training for another pair. The first pair in each chain received no training, but had to innovate gestures for the events. The events participants communicated about were presented as pairs of English sentences, containing one or both of the arguments Hannah and Sarah. Sentence pairs were either same-agent (where the agent in both sentences is the same), or different-agent (where the agent is different in sentence 1 and 2). Sentence pairs were presented within blocks of 4, with participants switching director/matcher roles after each block, and each block presenting either same-agent or differentagent sentence pairs. Within a block, each sentence pair demonstrated one of four verb types: verbs of motion (no endpoint), verbs of motion (endpoint), verbs of physical transfer, and verbs of non-physical transfer. The block structure presents a 'discourse' over which participants must repeatedly refer to Hannah and Sarah and differentiate between them to successfully communicate.

Gestures were coded on several parameters: type of gesture for agent and verb, location of agent and verb gesture, and path of verb gesture. Our results showed that the use of these parameters to disambiguate arguments in target sentences can be grouped into three main strategies: 1) a lexical strategy, in which participants vary the type of gesture used to denote the agent, 2) a body strategy in which body orientation (mapping onto verb location and path) signal the agent, and 3) the indexing strategy, in which participants index locations in the space around them (varying agent location). Over generations in the experiment, participants converge on a particular strategy, where it becomes further systematised. The strategies participants use reflect iconic strategies that find analogues in natural sign languages (e.g. lexical signs, body shift, and agent marking). Participants innovate highly iconic strategies that are systematized over generations in transmission chains. The evolution of participants' gestures reflects changes in emerging sign languages, showing a development from body-centered movement paths to gesture paths abstracted away from the body (Padden et al., 2010), and support experimental findings demonstrating the evolution of linguistic structure through cultural evolutionary processes (Kirby et al., 2015).

References

- Cormier, K., Fenlon, J., & Schembri, A. (2015). Indicating verbs in British Sign Language favour motivated use of space. *Open Linguistics*, *1*, 684–707.
- Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and Communication in the Cultural Evolution of Linguistic Structure linguistic structure. *Cognition*, 141, 87–102.
- Kocab, A., Pyers, J., & Senghas, A. (2014). Referential Shift In Nicaraguan Sign Language: A Transition From Lexical To Spatial Devices. *Frontiers in Psychology*
- Liddell, S. K. (2003). *Grammar, gesture and meaning in American Sign Language*. Cambridge: Cambridge University Press.
- Lillo-Martin, D., & Meier, R. P. (2011). On the linguistic status of "agreement" in sign languages. *Theoretical Linguistics*, 37(3–4), 95–141.
- Padden, C., Meir, I., Aronoff, M., & Sandler, W. (2010). The grammar of space in two new sign languages. Sign Languages: A Cambridge Language Survey. Cambridge University Press, Cambridge, UK, 570–592.

EVOLUTIONARY STABILITY OF LINGUISTIC POLITENESS AND THE POLITENESS EQUILIBRIUM PRINCIPLE

Roland Mühlenbernd¹, Przemysław Żywiczyński*², and Sławomir Wacewicz²

*Corresponding Author: przemek@umk.pl ¹Department of Management, Ca'Foscari University, Venice, Italy ²Department of English Studies, Nicolaus Copernicus University, Torun, Poland

1. Introduction

Unlike many facets of language - phonology, syntax, semantics or even pragmatics - linguistic politeness (LP) has attracted little attention of evolutionarily minded researchers. We think that this lack of interest - apart from a few isolated attempts (van Rooy, 2003; Żywiczyński, 2012; Wacewicz et al., 2015; Pleyer & Pleyer, 2016) is not dictated by a peripheral status of LP for the description of language, and especially for language evolution. LP is a universal characteristic of languages (cf. Brown & Levinson, 1987) but its specific markers are subject to considerable cultural variation, a combination of features that makes it an interesting target for evolutionary modeling. Next, LP is first and foremost a set of interactional strategies, and hence naturally lends itself to rendering in game-theoretic terms (cf. Quinley, 2011). In this paper, we take a game-theoretic approach and make a case that LP can be subsumed under a more general explanatory principle: disalignment of interests. This is formally expressed as the Politeness Equilibrium Principle (PEP), whereby the more disalignment there is between the interests of Speaker and Hearer, the more LP Speaker needs to use to offset the imbalance. Furthermore, we present a game-theoretic model to show that the use of LP predicated on the PEP forms an evolutionary stable system.

2. The Politeness Equilibrium Principle

Possibly, the main problem of rendering LP into evolutionary terms concerns its operationalization. LP, although admittedly elusive, is in principle measurable; e.g. it can be stated in terms of speech indirectness (Leech, 1983; Brown & Levinson, 1987), which in turn closely correlates with utterance length (Östman, 1989). Furthermore, following Brown and Levinson (1987), we assume the rationality of a competent Speaker, who is able to strategically increase and decrease the amount of LP in her utterances. Finally, we suggest that the use of LP involves social costs: the more LP Speaker uses, the more cost she has to pay. This point is

in line with the conceptualization of LP as social commodity (Werkhofer, 2005), whereby using politeness incurs Speaker's payment in the currency of *face* (Brown & Levinson, 1987) or *respect* (Kaplan, 1999).

In the present work, we focus on one type of speech act: *request*. In this regard, we assume that Speaker follows the following principle:

Politeness Equilibrium Principle (PEP): The amount of linguistic politeness Speaker uses in a request corresponds negatively to the alignment of interests¹ between Speaker and Hearer: the less aligned the interests are in favor of Speaker, the more linguistic politeness she uses.

3. The PPL model

Based on a principle in animal communication² (Johnstone, 1997; Zahavi & Zahavi, 1997; Searcy & Nowicki, 2005), we developed a game-theoretic model (PEP model) that shows that the strategic use of LP according to the PEP forms an evolutionarily stable system (Maynard Smith, 1982), but only under the following conditions: i) the usage of LP involves substantial costs (of social nature), ii) Speaker is rational in that she wants to maximize her payoff, defined as the difference between her expected utility of getting the request granted by Hearer and the costs involved, and iii) the disalignment of interests (reflecting how big a favor is being requested) does not exceed a particular degree: there are requests that Hearer will never grant, no matter how much LP Speaker uses. This last point is in line with more general theoretical work on game-theoretic signaling models that shows that signaling systems are not evolutionary stable when the magnitude of disalignment between Speaker and Hearer is too large (Crawford & Sobel, 1982).

4. Outlook

The PEP model is only a first step towards understanding the evolutionary stability of LP, and further work is necessary to make the model more complete. Most importantly, this includes taking into account other key variables that determine the amount of LP used in a request, such as social distance, power, or cultural competence. Furthermore, while the current version of the PEP model focuses on strategic behavior of Speaker, a more complete model will also take the strategic behavior of Hearer into consideration.

¹Note: In Politeness theory the role of alignment of interests was already discussed as a factor for the choice of LP in requests, termed by Brown and Levinson (1987) as *Rate of Imposition*.

²This principle follows the pattern: 'the more an individual needs, the more intensive is its signal' and can e.g. be found in food begging of offspring of birds (cf. Searcy & Nowicki, 2005, Fig. 1.2).

References

- Brown, P., & Levinson, S. (1987). *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.
- Crawford, V. P., & Sobel, J. (1982). Strategic information transmission. *Econometrica*, 50, 1431–1451.
- Johnstone, R. A. (1997). The evolution of animal signals. In J. R. Krebs & N. B. Davies (Eds.), *Behavioural ecology* (pp. 155–78). Oxford: Blackwell.
- Kaplan, D. (1999). What is meaning? explorations in the theory of meaning as use. brief version draft 1. University of California, Los Angeles.
- Leech, G. (1983). The principles of pragmatics. London: Longman.
- Maynard Smith, J. (1982). *Evolution and the theory of games*. Cambridge University Press.
- Östman, J.-O. (1989). Testing iconicity: Sentence structure and politeness. *Belgian Journal of Linguistics*, *4*, 145–163.
- Pleyer, M., & Pleyer, M. (2016). The evolution of im/politeness. In S. Roberts, C. Cuskley, L. McCrohon, L. Barceló-Coblijn, O. Fehér, & T. Verhoef (Eds.), *The evolution of language: Proceedings of the 11th international conference (evolang11).*
- Quinley, J. (2011). Politeness and trust games. In D. Lassiter (Ed.), *Proceedings* of the 2011 essli student session (pp. 143–151). Stanford University.
- van Rooy, R. (2003). Being polite is a handicap: Towards a game theoretical analysis of polite linguistic behavior. In *Tark: Proceedings of the 9*th conference on theoretical aspects of rationality and knowledge (pp. 45–58). New York: ACM.
- Searcy, W., & Nowicki, S. (2005). *The evolution of animal communication: Reliability and deception in signaling systems.* Princeton/Oxford: Princeton University Press.
- Wacewicz, S., Żywiczyński, P., & McCrohon, L. (2015). Linguistic politeness from an ethological perspective: Theoretical questions and empirical issues. *Theoria et Historia Scientiarum*, *11*, 81–98.
- Werkhofer, K. (2005). Traditional and modern views: The social constitution and power of politeness. In R. Watts, S. Ide, & K. Ehlich (Eds.), *Politeness in language: Studies in its history, theory and practice* (pp. 155–199). Berlin: De Gruyter Mouton.
- Zahavi, A., & Zahavi, A. (1997). *The handicap principle: A missing piece of darwin's puzzle.* New York: Oxford University Press.
- Żywiczyński, P. (2012). An attempt at a proxemic description of politeness from the ethological-evolutionary perspective. *Theoria et Historia Scientiarum*, *9*, 71–98.

THE INFLUENCE OF SHARED VISUAL CONTEXT ON PERFORMANCE WITH NOVEL CODES IN A REFERENTIAL COMMUNICATION GAME

THOMAS F. MÜLLER^{*1}, JAMES WINTERS¹, TIFFANY MORISSEAU², and OLIVIER MORIN¹

^{*}Corresponding Author: tmueller@shh.mpg.de ¹Minds and Traditions Research Group, Max Planck Institute for the Science of Human History, Jena, Germany ²Central European University, Budapest, Hungary

It has been proposed that ostensive communication enables complex languages to evolve (Scott-Phillips, 2014). Successful ostensive communication, both verbal and non-verbal, must rely on a body of shared information that has been described as "common ground" (Clark, 1996). By virtue of knowing similar things, communicators are better able to infer each other's intentions. As a subset of the general phenomenon of common ground, *shared visual context* refers to the situation of two interlocutors building on visibly common information. The importance of shared visual context for referential communication has been extensively studied (e.g. Clark, Schreuder, & Buttrick, 1983; Hanna, Tanenhaus, & Trueswell, 2003), but its impact on language evolution remains an open question. Other studies (Winters, Kirby, & Smith, 2015; Tinits, Nölle, & Hartmann, 2017) have investigated the effects of different types of context on evolving languages, but not considered the *shared* aspect of this context. How important is shared visual context in the genesis of new languages through communication?

We conducted two laboratory experiments to answer the question of whether shared visual context between two interlocutors could help solving the task of establishing novel codes, and using them for communication. In both experiments, dyads of participants played a game in which one subject was tasked to choose the target colour out of an array of four possible answers, while the second subject tried to communicate the target, using black-and-white abstract symbols only. The symbols represented abstract shapes and objects that are not limited to one natural colour - e.g., a diamond, a piece of candy, a butterfly. Shared visual context was manipulated between dyads by minimizing or maximizing what portion of the array seen by the receiver was also seen by the sender: In the *absent visual context* condition, the senders only knew about the target colour, whereas they had full access to the three distractors in the *shared visual context* condition. Crucially, participants never received feedback about correct or wrong answers during the game, but could only tell their partner whether they understood the message instead. This, together with the non-

conventionality of the symbols, ensured they were not relying on mere memorization, but actively creating novel codes during communication.

Our predictions for both experiments were preregistered on the Open Science Framework. In experiment 1, we tested 26 pairs of participants using the basic paradigm outlined above. In experiment 2, we aimed to replicate the main results in a new sample of 48 participant pairs. Additionally, the procedure was adjusted to test whether codes established with shared visual context were more generalisable and would still be used if contexts changed. We did this by splitting the basic experiment in two halves, during which the reference space was limited to one half of the total colour space. Further changes were a refined symbol set and a perceptually controlled colour space.

The results of both experiments show, as predicted, that performance was better in pairs with shared visual context than in pairs with absent visual context (see Figure 1). This could be confirmed in mixed-effects models with maximal random effects structure. Furthermore, the models demonstrate that pairs make significant progress in accuracy over time in both experiments.



Additionally, we compared the symbol inventories in both conditions in experiment 1, as measured by the number of symbols used successfully according to a preregistered threshold of successful trials, and found that symbol inventories were larger in shared visual context pairs. Lastly, we compared the generalisability of the codes in experiment 2, measured by calculating the relative number of symbols reused in the different context of the second half of the experiment. We found that reuse was only functional with shared visual context, yet absent visual context pairs reused more symbols.

Participants successfully solved the task by establishing new codes, which vary wildly between dyads. From the results, we conclude that shared visual context has an important role during the emergence of novel codes, making them more successful and frequent. This contributes to our understanding of how communication can arise ostensively.

References

- Clark, H. H. (1996). *Using language*. New York, US: Cambridge University Press.
- Clark, H. H., Schreuder, R., & Buttrick, S. (1983). Common ground and the understanding of demonstrative reference. *Journal of Verbal Learning & Verbal Behavior*, 22(2), 245-258.
- Hanna, J. E., Tanenhaus, M. K., & Trueswell, J. C. (2003). The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language*, 49(1), 43-61.
- Scott-Phillips, T. C. (2014). Speaking our minds: Why human communication is different, and how language evolved to make it special. Palgrave MacMillan.
- Tinits, P., Nölle, J., & Hartmann, S. (2017). Usage context influences the evolution of overspecification in iterated learning. *Journal of Language Evolution*, 2(2), 148-159.
- Winters, J., Kirby, S., & Smith, K. (2015). Languages adapt to their contextual niche. *Language and Cognition*, 7(3), 415-449.

A DOMESTICATED CODE: ON THE EMERGENCE OF THE OSCILLATORY BASIS OF PHRASE STRUCTURE

ELLIOT MURPHY*1,2

*Corresponding Author: elliot.murphy.13@ucl.ac.uk
¹Division of Psychology and Language Sciences, University College London, England
²Department of Psychology, University of Westminster, England

The search for the 'neural code' in animal neurophysiology has recently seen a marked transition from the analysis of individual spike timings to larger patterns of synchronisation, and it will be argued that the study of language should readily embrace these systems-level developments in the field. An emerging consensus in linguistics is that the language system needs to exploit two memory workspaces in order to generate and maintain/interpret non-local syntactic relations (e.g. Adger, 2017). One workspace is used to construct syntactic phrases, and the other is used to maintain these units. The findings of a number of MEG studies will be discussed in order to suggest that an initial phase-amplitude coupling θ - γ code constitutes the first workspace, and a subsequently embedded δ - θ code constitutes the second. It will be shown that this account goes some way to explain why left-inferior frontal regions seem to be implicated in semantic composition operations.

If the particular δ - θ interactions documented here are indeed humanspecific, this could derive the apparent uniqueness of the dual workspace of human syntax. δ appears vital for language comprehension, but from a neurocomputational perspective, cross-frequency coupling involving δ may also play a vital role in the online construction of phrases, instead of 'passively' entraining to phrasal units (Ding et al., 2016). Although cross-frequency coupling is most commonly found between θ and γ , recent work suggests that θ entrains to δ in the human brain in a uniquely diverse way (Maris et al., 2016). Given experimental findings that these slow rhythms entrain to sentential and phrasal structures (Mai et al., 2016; Bradley & Iverson, 2017), the discovery of humanspecific forms of δ - θ phase-amplitude coupling is potentially of great significance; in particular when interpreted alongside findings that non-human primate call systems exhibit limited cross-frequency couplings, as shown in Murphy (2016a-c, Forthcoming a,b). Grimaldi (2017: 17) summarises that 'Murphy [(2016b)] proposes a revolutionary theory of neurolinguistics ... that

nested oscillations execute elementary linguistic computations', adding that this theory 'goes considerably beyond existing models'. This model will be refined, taking into account recent findings concerning, for instance, travelling oscillations.

It will further be argued that this phase code can be explained through reference to the domestication syndrome and the 'globularity' hypothesis (Benítez-Burraco et al., 2017), such that the human braincase permits a reduction in 'spatial inequalities' (Salami et al., 2003) between cortical and subcortical regions documented in less globular brains. The MEG data presented will be shown to be in line with this position.

In addition, a critique of recent neuroethological work is presented. For instance, Kikuchi et al. (2017) claim to find no differences in the cross-frequency coupling profiles of humans and monkeys, however they only examine coupling between low frequencies and γ , and not between low frequencies such as δ , θ and β . Upon its emergence, the human-specific oscillatory profile would have interfaced with more ancient systems shared with other primates, in turn influencing and modifying the internal organisation of such systems through a process of 'reciprocal causation' (Walsh 2015). As a result, the emergence of the presently proposed neural code would likely have re-shaped the representations of the conceptual systems it interfaced with. It is argued that this can explain why primate call units appear to be lacking in the diverse array of language-specific representations. More specifically, it is argued that domestication can be used as a model system for the Extended Evolutionary Synthesis (Zeder, 2017) and can enhance our understanding of reciprocal causation and niche-construction, which in turn allows neurolinguistics to make more direct contact with the life sciences.

Lastly, the potential for cross-frequency coupling to derive the computational complexity of hierarchically organised phrase structures is explored in order to construct a set of working hypotheses concerning the role of particular neural oscillations in linguistic computation. Processes once deemed highly domain specific – from grid cells to the theta-gamma code for working memory – are increasingly being shown to execute generic computations operating over items stored across distinct representational domains (Constantinescu et al., 2016; Kriegeskorte & Storrs, 2016; Lisman & Jensen, 2013), and the consequences of this for the language sciences are presented. It will be argued that a number of recent studies claiming to make substantial insights into the oscillatory nature of language have engaged primarily in data redescription rather than theory-construction, and this discussion will be used to motivate a neurobiologically feasible model of syntactic computation which brings with it direct consequences for more traditional, cartographic models.

Acknowledgements

This work was supported by an Economic and Social Research Council scholarship (1474910).

References

Adger, D. (2017). A memory architecture for merge. Ms. Queen Mary University.

- Benítez-Burraco, A., Di Pietro, L., Barba, M., & Lattanzi, W. (2017). Schizophrenia and human self-domestication: an evolutionary linguistics approach. *Brain, Behavior and Evolution*, 89(3), 162-184.
- Bradley, H., & Iverson, P. (2017). How do infants process speech in the mind? An EEG study of neural entrainment and the Frequency Following Response. Poster presented at the UCL MRes/MSc Language Sciences Poster Session.
- Constantinescu, A. O., O'Reilly, J. X., & Behrens, T. E. J. (2016). Organizing conceptual knowledge in humans with a gridlike code. *Science*, 352(6292), 1464-1468.
- Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience*, 19, 158–164.
- Kikuchi, Y., Attaheri, A., Wilson, B., Rhone, A. E., Nourski, K. V., Gander, P. E., Kovach, C. K., Karasaki, H., Griffiths, T. D., Howard III, M. A., & Petkov, C. I. (2017). Sequence learning modulates neural responses and oscillatory coupling in human and monkey auditory cortex. *PLOS Biology*, 15(4), e2000219.
- Kriegeskorte, N., & Storrs, K. R. (2016). Grid cells for conceptual space? *Neuron*, 92, 280-284.
- Lisman, J. E., & Jensen, O. (2013). The theta-gamma neural code. *Neuron*, 77, 1002-1016.
- Mai, G., Minett, J. W., Wang, W. S.-Y (2016) Delta, theta, beta, and gamma brain oscillations index levels of auditory sentence processing. *NeuroImage*, 113, 516-528.
- Maris, E., Fries, P., & van Ede, F. (2016). Diverse phase relations among neuronal rhythms and their potential function. *Trends in Neurosciences*, 39(2), 86-99.
- Murphy, E. 2016a. Evolutionary monkey oscillomics: Generating linking hypotheses from preserved brain rhythms. *Theoretical Linguistics*, 42(1-2), 117-137.
- Murphy, E. 2016b. The human oscillome and its explanatory potential. *Biolinguistics*, *10*, 6-20.
- Murphy, E. 2016c. A theta-gamma neural code for feature set composition with phase-entrained delta nestings. UCL Working Papers in Linguistics, 28, 1-23.
- Murphy, E. Forthcoming a. *Principles of Language Evolution*. [Textbook]. London: UCL Press.
- Murphy, E. Forthcoming b. Implications of travelling weakly coupled oscillators for the cortical language circuit. UCL Working Papers in Linguistics.

- Salami, M., Itami, C., Tsumoto, T., Kimura, F. (2003). Change of conduction velocity by regional myelination yields constant latency irrespective of distance between thalamus and cortex. *PNAS*, 100, 6174-6179.
- Walsh, D. M. (2015). *Organisms, Agency, and Evolution*. Cambridge: Cambridge University Press.
- Zeder, M. A. (2017). Domestication as a model system for the extended evolutionary synthesis. *Interface Focus*, 7, 20160133.

SEMANTIC BLEACHING NOT OBSERVED IN SYNCHRONIC TEST

Dillon Niederhut*1

*Corresponding Author: dillon.niederhut@gmail.com ¹Enthought Inc., Austin, United States of America

It is well understood that the semantic content of words changes over time, but precisely how and why this happens is still unknown. Here, we test for synchronic evidence of semantic bleaching in a corpus of English language collected in 2016. We find no evidence of long-term reduction in the semantic value of words, although this may not be true when considered over shorter periods of time.

1. Introduction

SEMANTIC BLEACHING is a well-described phenomenon where the specificity of a word decreases with use. To take one common example, the term *awesome* was once reserved for the Judeo-Christian deity, but is now used to describe everything from toast¹ to the Transformer movie franchise². Another way to state this observation is that a word which first refers one thing (which is presumably the case for all words) can be generalized over time to also refer to other, related things. In the example given here, *awesome* has been extended to include all sorts of things which supposedly inspire feelings of religious devotion.

Semantic bleaching may be part of a more general process whereby words expand and shift their meanings in order to maximize some definition of communicative optimum. Piantodosi, Tily, and Gibson argue that the driving force is the production of speech, and that it is optimal to find new uses for phrases with short orthographic length in order to reduce the overall number of graphemes or syllables needed to convey any particular idea (Piantadosi, Tily, & Gibson, 2011). Recent work by Xu has also incorporated the cost of interpretation of words, focusing on reducing the ambiguity during the process of assigning labels to objects. In one semantic domain, the historical shift in word meanings over time approaches a Pareto frontier balancing the cost of production with the cost of interpretation for assigning existing words to new kinds of containers (Xu, Regier, & Malt, 2016).

¹http://www.bonappetit.com/restaurants-travel/article/ how-to-make-perfect-toast

²http://kotaku.com/

leave-michael-bay-alone-transformers-is-awesome-1596887614
Here, we conduct a synchronic test for historical evidence of semantic bleaching in an English corpus collected in 2016. Under the assumption the rate of bleaching per word tends to be positive, i.e. that it outpaces the rate of fossilization, then it follows that in aggregate older words should have less specific meaning than newer words. This must be certainly be true in the narrow sense, at least for the coinage of new terms, but it remains to be seen whether this relationship holds over historic time.

2. Methods

The date of first written appearance of a large number of English words was acquired from Merriam Webster's recently published "Words by First Known Date" (Merriam Webster, 2018). This generated dates for approximately 10,000 terms, with a roughly even distribution over the last 100 years, and appearances of terms becoming less specific further than the 1800s (Fig. 1)



Figure 1. Distribution of estimates of the year of first appearance for 1,000 English words; each bin has a width of ten years

Term frequency was acquired from redicorpus dataset described in Niederhut

Niederhut (2017). Briefly, this corpus includes approximately 4.3 million communication events sampled from the AskReddit subgroup of reddit.com over 22 weeks in 2016. We were able to discover daily use frequencies and communication contexts for approximately 2500 of the terms from Merriam Webster after applying the Porter Stemmer (Porter, 1980). While redicorpus contains terms of all kinds (including function words), the matches with the Merriam Webster data are largely nouns and verbs like *unfriend*, *mantissa*, and *carotid sinus*.

Zipf statistics, a measure of semantic value, were calculated according to the formula from (Niederhut, 2016). The intuition is that words like *multicollinear* feel like they contain more information than words like *my*, because the former provides evidence of a rather specific context (a problematic linear model) whereas the latter does not. The belief that an utterance comes from a specific linguistic context creates an expectation that other context-appropriate terms like *variance inflation factor* will also be observed. The Zipf statistic quantifies the magnitude of this expectation by comparing the probability distribution of words conditioned on one context against the distribution of words across all contexts.

We were able to calculate Zipf statistics for about 300 of the Merriam Webster terms. This number is remarkably smaller due to the computational cost of calculating these statistics. Each test statistic takes roughly 90 minutes to process, largely due to the I/O cost of reading large amounts of text data. The sample thus represents roughly 435 CPU hours, or eighteen days, of work.

	year	frequency	Zipf statistic
Min.	1799	9.3e-09	-0.794
1st Qu.	1879	2.8e-08	1.495
Median	1926	1.3e-07	2.534
Mean	1919	1.6e-07	7.354
3rd Qu.	1954	6.5e-07	4.653
Max.	2009	8.3e-04	233.577
Ν	10191	2594	290

Table 1. Summary statistics for year of first appearance, frequency of use in 2016, and estimated Zipf statistic

A linear model was run regressing the Zipf statistics on the year of first appearance of each word. To help control for the effect of any outliers, the model was rerun using the robust linear modeling package for R (Wang et al., 2014). To test for evidence of semantic bleaching over short time scales, a third model was run including only data since the 1990s. To control for the possibility that the outcome might be determined by some peculiarity of the Zipf statistic, we also tested for a relationship between the age of a term and its daily moment³.

³The daily moment of a term is it's average daily uses divided by the standard deviation in its daily

Data were collected in Python 3.5.4 on Ubuntu Server 16.0.4, and were analyzed with Revolution R Open⁴ based on CRAN release v. 3.2.3, "Wooden Christmas Tree", (R Core Team, 2015). Tables were produced with xtable, and figures were produced with ggplot2 (Dahl, 2014; Wickham, 2009). Data and R files to reproduce this analysis along with its tables and figures are available at https://github.com/deniederhut/semantic-bleaching-not-observed-in-synchronic-test.

3. Results

Table 2.	Model statistics from an OLS regressing Z	Lipf statis-
tic on yea	ar of appearance.	

	Estimate	Std. Error	t value	Pr(> t)				
(Intercept)	-25.42449	45.23668	-0.562	0.575				
year	0.01706	0.02353	0.725	0.469				
Residual standard error: 21.65 on 288 degrees of freedom								
Multiple R-squared: 0.001821, Adjusted: -0.001645								
F-statistic: 0.5255 on 1 and 288 DF, p-value: 0.4691								

We find no evidence that the Zipf statistic of a word is related to how long that word has been in use. The first linear model assigns a coefficient of 0.03 to the year term, which is not significantly different than zero at p = 0.20 (Table 2). The R^2 for the model, in both the corrected and uncorrected estimates, is less than 1%.

The robust model, which ignores roughly 40 of the very large Zipf values, produces similar results, with a coefficient for the year term that is less than 0.01, and not significantly different from zero with p = 0.38 (see Fig. 2). The R^2 for the model, in both the corrected and uncorrected estimates, is less than 1%.

		<u> </u>		D (11)				
	Estimate	Std. Error	t value	$\Pr(> t)$				
(Intercept)	-154.23339	116.92727	-1.319	0.194				
year	0.07844	0.05865	1.337	0.188				
Robust residual standard error: 1.766								
Multiple R-squared: 0.04715, Adjusted R-squared: 0.02499								

Table 3. Model statistics from a robust model regressing Zipf statistics on years after 1990.

The recent model, which includes only the last twenty years in the data set, offers some mild evidence toward a short-term effect of time on semantic value,

use, and can be interpreted as a measure of generality. Common words like *deny* have large moments (here, cerca 3.2), while uncommon words like *bantamweight* have small moments (here, c. 0.08)

⁴https://mran.revolutionanalytics.com/open/



Figure 2. Relationship between Zipf statistics and year of appearance, with outliers (as determined by robust algorithm) removed, with superimposed trend line from robust model.

with a positive effect coefficient of 0.08 and p < 0.20 (Table 3). The R^2 for the model, in both the corrected and uncorrected estimates, is roughly 4%.

	Estimate	Std. Error	t value	Pr(> t)				
(Intercept)	0.0008698	1.1729085	0.001	0.999				
year	0.0002458	0.0006102	0.403	0.687				
Residual standard error: 0.5612 on 288 degrees of freedom								
Multiple R-squared: 0.0005633, Adjusted: -0.002907								
F-statistic: 0.1623 on 1 and 288 DF, p-value: 0.6873								

Table 4. Model statistics from an OLS regressing daily moment on year of appearance.

The comparison against the daily moment produces the same null result as the overall and robust tests against the Zipf statistic, with no significant relationship between the year of first appearance of a term and the generality with which it is used in natural language production.

4. Discussion

It is interesting that we find no relationship between the semantic value of a word, as measured by the Zipf statistic, and how long that word has been in use. Prima facie, the oldest words in a language should have had more time to take on additional meanings, and to have had their specificity diluted through use in metaphor and other poetic devices. One interpretation of this finding is that bleaching takes different trajectories for different words, and that these changes will only be visible in diachronic tests.

Another possibility is that semantic bleaching, where the specificity of a word decreases over time, is balanced by a force that removes less-used senses of the word in order to reduce ambiguity in its meaning. This may be a passive process, where some semantic interpretations fall out of use simply because their referents do, like the use of *wire* to describe information transfer over telegraph.

A final interpretation is that Merriam Webster has chosen particularly interesting words about which to publish the year of appearance online. Based on prior work, we would expect a selection of words randomly sampled from human usage to have words that appear much more frequently. The words in this sample have an average proportion of 2.2e - 06, and a median of 1.3e - 07 (see Table 1). The presence of words like *wiki* and *a tempo* may also explain the unusually high Zipf statistics observed for these data.



Figure 3. Relationship between Zipf statistics and year of appearance since the 1990s, with outliers (as determined by robust algorithm) removed, with superimposed trend line from robust model.

Given the findings from the third model, it remains possible that words bleach very quickly over a short duration, and then reach some kind of equilibrium in the population whose stable point is governed by other forces, likely including the size of the speech community (Fig. 3). However, the magnitude of the "short term bleaching" effect that we have measured is small enough that we do not feel comfortable arguing that it provides a plausible alternative to the definite null effect over longer periods of time without corroborating evidence of its existence.

References

Dahl, D. (2014). xtable: Export tables to latex or html. (R package version 1.7-4)

- Merriam Webster. (2018). Words by first known use date. https://www.merriam-webster.com/words-by-first-known-date.
- Niederhut, D. (2016). Quantifying the semantic value of words. In S. Roberts, C. Clusky, L. McCrohon, L. Barcelo-Coblijn, O. Feher, & T. Verhoef (Eds.), *The evolution of language: Proceedings of the 11th international conference.* Hackensack: World Scientific.
- Niederhut, D. (2017). *Performance approaches to semantics in human language*. Unpublished doctoral dissertation, University of California at Berkeley.
- Piantadosi, S., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Science*, 108, 3526-3529.
- Porter, M. (1980). An algorithm for suffix stripping. Program, 14, 130-137.
- R Core Team. (2015). *R: A language and environment for statistical computing.* Vienna, Austria.
- Wang, J., Zamar, R., Marazzi, A., Yohai, V., Salibian-Barrera, M., Marrona, R., Zivot, E., Rocke, D., Martin, D., Maechler, M., & Konis, K. (2014). *robust: Port of the s+ "robust" library.* (R package version 0.4-16)
- Wickham, H. (2009). ggplot2: elegant graphics for data analysis. Springer New York.
- Xu, Y., Regier, T., & Malt, B. (2016). Historical semantic chaining and efficient communication: the case of container names. *Cognitive Science*, 40, 2081-2094.

COMPARATIVE MORPHOLOGY OF THE LARYNGEAL MUSCLES IN HYLOBATIDS USING A HIGH-RESOLUTION MRI

TAKESHI NISHIMURA^{1,*}, YOSHITAKA NOMURA², HIROHIKO IMAI³, TETSUYA MATSUDA³

*Corresponding Author: nishimura.takeshi.2r@kyoto-u.ac.jp ¹Primate Research Institute, Kyoto University, Inuyama, Japan ²Graduate School of Science, Kyoto University, Kyoto, Japan ³Graduate School of Informatics, Kyoto University, Kyoto, Japan

Hylobatids are lesser apes inhabiting the canopy of the tropical forest in Southeast Asia, and they are usually classified into the four independent genera, i.e., Hylobates, Hoolock, Nomascus, Symphalangus. They produce characteristic calls, which are referred to as "song", which comprise a sequence of loud and pure-tone-like voices changing melodiously in fundamental frequency (f0), i.e., pitch. The physiological mechanisms of animal vocalization are often examined by the acoustics of voices recorded in a helium-enriched atmosphere: so-called 'helium voices' (e.g., Nowicki, 1987). The helium voices demonstrated that song voices in gibbons are also produced based on the source-filter independence, like in human speech (Koda et al., 2012). Gibbons use the sound source generated by vibration of the bilateral vocal folds (VFs), to amplify f0 and to suppress the higher harmonics in a filter of the supralaryngeal vocal tract (SVT). This means that modifications of the pitch location principally depend on rapid and flexible modifications of VF elasticity in gibbons. The VF is connected between the anterior commissure of thyroid anteriorly and bilateral arytenoids on the cricoid posteriorly. The VF elasticity is modified by contractions of the vocal muscles that is one of the major components of VF and by varied movements of laryngeal cartilages which extends the VFs in humans. Thus, some anatomical adaptations are expected for large modifications in VF elasticity in gibbons.

We here use the MRI scans of the embalmed specimens of extracted larynges, to examine the morphological variation of laryngeal architecture in gibbons and other non-human primates. The specimens from the four genera of hylobatids, and chimpanzees, orangutans, macaques, were scanned by a high-resolution MRI scanner (7T MRI BioSpec 70/20 USR, circular polarized transceiver volume coil 72-mm bore, T9562), with a special resolution of 54.7-97.7µm³ and FOV of 28-50mm², depending on specimen's size. Three-dimensional configuration of each

laryngeal skeleton and internal laryngeal muscle was reconstructed using 3D visualizing software of FEI Amira 6. We also examined the external laryngeal muscles by anatomical dissection.

Some features characteristic and common to hylobatids are found in MRI scans: the vocal muscle is very thin in medio-lateral direction; the first tracheal cartilage is located inside the cricoid cartilage, e.g., intracricoid position; the inner and outer layers of the cricothyroid muscle extend upward and is attached to the middle to superior surface area of thyroid lamina, respectively. By contrast, the other primates have the features that are also found in humans: the thick vocal muscle; the first tracheal cartilage is located inferior to the cricoid; the cricothyroid is attached to the inferior edge or area of the lamina and the inferior horn of thyroid.

The features in gibbons are suggested to underlie their high-pitched and flexible phonation. Their modifications in f0 principally depends on the modifications of the spatial relationship of the laryngeal cartilages, e.g., the rocking of the thyroid against the cricoid-arytenoid cartilages rather than of the vocal muscle elasticity. Vocal muscle is very tiny in gibbons, indicating its contraction makes limited contributions to the modifications in the VF elasticity. On the other hand, their long crico-thyroid muscle makes the thyroid rock largely against the cricoid. Such a large rocking requires less efforts for a rapid and repeated modifications in VF elasticity, compared to contractions of the vocal muscle. Such derived features are available for rapid and repeated modifications in pitch in gibbons' song, which is different from mild vocalizations in the other primates examined here.

Humans have anatomical specifications of a two-tube configuration of SVT, which is adapted for rapid and flexible modifications of formants. Gibbons also have another anatomical specification for their melodious singing, but alternatively such specification is not advantageous to mild vocalizations like speech. While humans and gibbons share a common voice physiology, anatomical variation arose adapted for different physiological manipulations of vocalizations varied in primates including humans. Increasing knowledge of the anatomy and physiology in non-human primates promises better understanding of primate origins and of the evolutionary history of physical faculties in human speech.

Acknowledgements

This study is financially supported in part by JSPS KAKENHI (#16H04848, TN). We thank Japan Monkey Centre and Yokohama "Zooasia" for a permission to use specimens.

- Koda, H., Nishimura, T., Tokuda, I. T., Oyakawa, C., Nihonmatsu, T., & Masataka, N. (2012). Soprano singing in gibbons. *American Journal of Physical Anthropology*, 149, 347-355.
- Nowicki, S. (1987). Vocal tract resonances in oscine bird sound production: evidence from birdsongs in a helium atmosphere. *Nature*, *325*, 53-55.

ENVIRONMENTAL AND SOCIAL FACTORS MOTIVATE THE EMERGENCE OF SYSTEMATIC CATEGORIES AND SIGNS

Jonas Nölle*1,3, Marlene Staib², Riccardo Fusaroli^{2,3}, and Kristian Tylén²

*Corresponding Author: j.nolle@sms.ed.ac.uk
¹Centre for Language Evolution, University of Edinburgh, UK
²Center for Semiotics, Aarhus University, Denmark
³The Interacting Minds Centre, Aarhus University, Denmark

While arbitrariness has long been considered a hallmark of human language, there have been increasing discussions about non-arbitrary relationships between form and meaning such as iconicity and systematicity (Dingemanse et al., 2015). We argue that these phenomena are not just two facets of non-arbitrariness, but serve orthogonal functions in the scaffolding of an efficient communication system. Iconicity is usually associated with learning and bootstrapping (Imai & Kita, 2014), while systematicity has been shown to emerge in the lab when iconicity is inhibited (Roberts et al., 2015; Verhoef et al., 2016) or when participants were provided with pre-established combinatorial categories (Theisen et al., 2010).

Building on this work, we tested in a series of dyadic silent gesture experiments whether systematicity is functionally adaptive and could also emerge in competition to iconicity in response to particular environmental and social factors. We hypothesized that structure and openness of the environment as well as working memory constrains (Christiansen & Chater, 2016) could affect the degree of systematicity as well as which referential features become systematized. In two experiments, participants had to communicate stimuli depicting stylized characters that afforded both iconic and systematic gestures (see Fig. 1A). Experiment 1 tested a) whether functional adaptivity of a given trait, based on its distribution in the referential environment (consisting of 14 stimuli drawn from a set of 24) and thus its discriminative value, would affect the likelihood of this trait being systematized and b) the impact of an expanding referential environment. In a 2×2 design we varied the distribution of PET vs. GENDER (7:7 vs. 10:4) and the openness of the environment (an expanding meaning space in the *open* vs. constant meanings across trials in the *closed* condition). Logistic mixed regression models (see supplementary materials) indicated that functionally more adaptive traits were indeed more likely to be systematized (p < .01) while openness of the environment had no significant effect (see Fig. 1B). Debriefing revealed that participants in the open condition frequently realized that the changing competitor images never had to be signed, which narrowed down the search space rather than enlarging it.



Figure 1. A Stimuli possessed various salient idiosyncratic features affording iconic descriptions (e.g., hat, scarf) as well as categorical traits (e.g., FEMALE, CLERICAL, PET) that could be fully systematized. B Systematicity by functional adaptivity and referent environment (exp. 1) C Mean proportion of systematic gestures by trial number (exp. 2). See supplementary materials for statistics.

Experiment 2 was designed to further explore the roles of informational bottlenecks. We improved the "openness" manipulation by increasing the number of stimuli (n=32) and referents per trial (n=16) and randomly sampling both targets and competitors for each trial in the open condition, while the set remained the same across trials in the *closed* condition. In addition, we added "displacement" as a second factor. As in experiment 1, in the co-present condition, participant pairs could gesture while seeing the stimuli, whereas dyads in the *displaced* condition had to wait 3 seconds after the stimuli disappeared before they could gesture-effectively displacing the communicative from the referential context and simulating communication about absent entities. Our results indicate that displacement increased the tendency to systematize overall, while openness of the environment affected the temporal development of systematicity (see Fig. 1C). In the open condition systematicity kept increasing until the end of the experiment, while there was a decline in the closed condition that resembled the slope of both the closed and "pseudo"-open condition in experiment 1, suggesting that a true open environment can override the tendency to fall back on simple iconic signs as communication becomes more efficient over time (Kirby et al., 2015).

Taken together, our results indicate that, given certain environmental affordances, systematicity can emerge in a novel communication system at the interaction level even in the presence of competing iconic solutions. In our experiment both systematic categories and signs emerged in response to their functional adaptivity to the environment and working memory constraints of the communicative situation (openness of the referential context, displacement). This could lead to variation that is amplified over cultural transmission (Kirby et al., 2015). Lastly, our study provides the first experimental semiotics study investigating the effect of displacement, a core property of language (Hockett, 1960) that has curiously been neglected in experimental studies so far.

Acknowledgements

We would like to thank the Cognition and Behavior Lab at Aarhus BSS for facilities and equipment as well as Sergio Gonzales De La Higuera Rojo and Caroline Kildahl for coding of the video material. This work was funded by a Seed Grant of the Interactive Minds Centre as well as a funding from the School of Communication and Culture, Aarhus University.

- Christiansen, M. H., & Chater, N. (2016). The Now-or-Never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*, 39.
- Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, Iconicity, and Systematicity in Language. *Trends in Cognitive Sciences*, 19(10), 603–615.
- Hockett, C. F. (1960). The Origin of Speech. Scientific American, 203(3), 88-97.
- Imai, M., & Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Phil. Trans. R. Soc. B*, *369*(1651), 20130298.
- Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, 141, 87–102.
- Roberts, G., Lewandowski, J., & Galantucci, B. (2015). How communication changes when we cannot mime the world: Experimental evidence for the effect of iconicity on combinatoriality. *Cognition*, *141*, 52–66.
- Theisen, C. A., Oberlander, J., & Kirby, S. (2010). Systematicity and arbitrariness in novel communication systems. *Interaction Studies*, *11*(1), 14–32.
- Verhoef, T., Kirby, S., & de Boer, B. (2016). Iconicity and the Emergence of Combinatorial Structure in Language. *Cognitive Science*, 40(8), 1969–1994.

THE EVOLUTIONARY CAUSES OF IRREGULARITY: EVIDENCE FOR AN IRREGULARIZATION BIAS IN MORPHOLOGICAL LEARNING

TIMOTHY J. O'DONNELL^{*1} and KENNY SMITH²

^{*}Corresponding Author: timothy.odonnell@mcgill.ca ¹Department of Linguistics, McGill University, Québec ²Centre for Language Evolution, University of Edinburgh, UK

While inflectional morphology is broadly rule-governed, many inflectional paradigms admit some exceptions (e.g., the past tense of "go" is not "goed" but "went"). Regularity in form-meaning mapping permits generalization and facilitates learning, and regularity has been shown to emerge through language transmission (e.g. Kirby et al., 2008, 2015).

The presence of irregulars is more puzzling. One possibility is that irregulars are simply a by-product of processes which destroy regularity (e.g., minimisation of effort in production), which survive language transmission due to their high frequency (Kirby, 2001; Lieberman et al., 2007). Here we show that irregularity in the right place in a morphological system actually *facilitates* the learning of the productive parts of the system; therefore, irregularity might be favoured during language transmission, rather than simply being a by-product of other processes.

We build on a recent computational model (O'Donnell, 2015) which shows that high-frequency irregulars facilitate the learning of productive regular rules. This model treats learning as an inference problem, where learners infer the productivity of morphological processes, balancing a tradeoff between computation and storage. High-frequency items tend to be stored as wholes, rather than handled compositionally; productive computation is signaled by morphological processes which apply across a large number of low-frequency forms. Since there is pressure to store high-frequency regular forms, they detract from the productivity of regular inflectional processes. When high-frequency forms are instead irregular, regular rules generalize more easily.

We provide experimental evidence for this *irregularization bias*. 46 adult participants learned 48 novel inflected words, organized into six disjoint paradigms of eight words each (see Table 1). In two *Fully Regular* sets, all stems were inflected with a single regular suffix and occurred with equal frequency during training. The remaining four sets had eight stems that occurred with non-uniform frequency. In each set, seven stems occurred with a regular suffix while a single item took an irregular suffix. These sets differed in whether the the most frequent

Table 1. Example word sets from the critical conditions, giving frequency of exposure in training for each word form and an example definition. Irregulars are highlighted in bold/red. The Fully Regular sets (not shown) have 8 stems, each occurring with frequency 12 and inflecting with a single regular suffix.

Frequ	ient Irregular Set	Freque	ent Regular Set
Frequency	Word	Frequency	Word
48	shrunefench	48	fibimort
17	chunwobane	17	dwilnebort
10	yadnossane	10	spolakort
6	raldane	6	moyport
6	shrumane	6	chervort
4	benthane	4	glocktonort
3	flovenane	3	quideport
3	heespane	3	dagzomuth
Example de	finition: "shrunefench,	Example defin	nition: "fibimort, the
the tool used	to shrunef"	place where th	ey make fibim"

item was irregular (*Frequent Irregular* sets) or regular (*Frequent Regular* sets). Participants were trained and tested over two days. Each day involved a training phase (auditory exposure to inflected forms plus definitions), followed by a test (participants were presented with an inflected word and asked if they had encountered it during training); these tests included trials which tested generalization of affixes, and trials which tested memorization of specific inflected forms.

Following the predictions of O'Donnell (2015), participants were more likely to generalize the regular affix from the Frequent Irregular set than the regular affix from the Frequent Regular set. This was not merely due to the higher frequency of regular suffixes from the Frequent Regular sets, since Fully Regular suffixes patterned with the Frequent Irregular suffixes (i.e. were highly productive, despite having frequency similar to the regular suffixes from the Frequent Regular sets). We also observed a storage advantage for irregulars: high-frequency irregulars were more likely to be accepted than frequency-matched high-frequency regulars.

These results demonstrate that high-frequency irregulars facilitate the generalization of regular rules for human language learners, or conversely that highfrequency regulars inhibit the generalization of a regular rule. This shows that systems in which high-frequency items are irregular have learnability advantages. In current work we are investigating whether this results in such systems emerging through language transmission; if so, irregularity might be explained by the same mechanisms already invoked to explain the evolution of regularity.

Acknowledgements

This project has received funding from the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme (grant agreement 681942).

- Kirby, S. (2001). Spontaneous evolution of linguistic structure: an iterated learning model of the emergence of regularity and irregularity. *IEEE Transactions on Evolutionary Computation*, 5(2), 102–110.
- Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: an experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences, USA*, 105, 10681-10686.
- Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, 141, 87-102.
- Lieberman, E., Michel, J.-B., Jackson, J., Tang, T., & Nowak, M. A. (2007). Quantifying the evolutionary dynamics of language. *Nature*, 449, 713–716.
- O'Donnell, T. J. (2015). *Productivity and reuse in language: A theory of linguistic computation and storage*. Cambridge, Massachusetts and London, England: The MIT Press.

EVOLUTION OF HOMOPHONES AND SYNTACTIC CATEGORIES NOUN AND VERB

Mieko Ogura^{*1} & William S-Y. Wang²

*Corresponding Author: ogura-m@tsurumi-u.ac.jp ¹Linguistics Laboratory, Tsurumi University, Yokohama, Japan ²Language and Cognitive Sciences, Hong Kong Polytechnic University, Hong Kong, China

We examine the neural substrates of comprehension of the bisyllabic homophones in English and Japanese. The evolution of homophones is a result of interaction between the speaker's production and the listener's perception, and the cortical representation of speech does not merely reflect the external acoustic environment. We further demonstrate that noun and verb categories are represented in different neural substrates in English, while both categories are processed in the same anatomical area in Japanese. We suggest that nouns and verbs are originally processed in the same anatomical area of the brain when the basic word order is SOV, and discourse organization is closely interwoven with syntactic organization as shown in Old English and Japanese. When the SOV order is changed to SVO through embedding, a strictly syntactic organization of the clause appears. The noun/verb distinction is clearly visible in the positions that nouns and verbs occupy in sentences, and it is represented in different neural substrates in the brain as shown in Present-day English.

1. Introduction

Zipf (1949) suggests the simultaneous minimization of the two opposing forces from listener and speaker for form and meaning associations. Zipf's law, which states that word frequencies decay as a power law of its rank, is the outcome of form-meaning associations adopted for complying with listener and speaker needs. Zipf's law implies one form to many meanings, i.e., polysemy and homophony. Polysemy and homophony are the necessary conditions for symbolic systems. Homophony is a desirable feature in communication system and likely results from ubiquitous pressure from efficient communication (Ferrer i Cancho & Solé 2003, Piantadosi et al. 2012, Wang 2011, Ogura forthcoming).

In this study, we examine the neural substrates of comprehension of homophones. We investigate the evolution of homophones as a result of

interaction between the speaker's production and the listener's perception. We further demonstrate that noun and verb categories are represented in different neural substrates in English, while both categories are processed in the same anatomical area in Japanese, and explore why such differences arose.

2. Neural Substrates of Bisyllabic Noun-Verb Homophones in English

Sereno (1986) conducts the analysis of bisyllabic noun-verb ambiguous pairs, i.e., words with both noun and verb instances. The 570 bisyllabic noun-verb ambiguous words in the Brown Corpus are analyzed in terms of location of the stressed syllable and frequency dominance of the grammatical category. Words are categorized as either forestressed, backstressed, or variably stressed. Dominance is characterized as noun-dominant, verb-dominant, or equidominant. A noun- or verb-dominant word is a word which occurs more than 50% of the time as a noun or verb, respectively. The vast majority (90%) of the ambiguous group includes words that maintain a constant stress pattern across syntactic function. They show consistent stress placement depending on their dominant (75%) than verb-dominant (22%) while backstressed stimuli are only marginally more noun-dominant (55%) than verb-dominant (43%).

Sereno & Jongman (1995) investigate systematic acoustic differences in bisyllabic noun-verb ambiguous pairs. 5 speakers read 16 bisyllabic words in both noun and verb contexts. The 16 bisyllabic homophones consist of four categories: 4 pairs which occur more frequently as forestressed nouns (favor, poison, practice, struggle); 4 pairs which occur more frequently as backstressed nouns (control, debate, dispute, report); 4 pairs which occur more frequently as forestressed verbs (handle, notice, rescue, welcome); and 4 pairs which occur more frequently as backstressed verbs (embrace, escape, neglect, reply). They find that stimuli that are more frequent as nouns in English show significantly different first/second syllable amplitude ratios than word stimuli that are more frequent as verbs. In other words, the amplitude of the first syllable of word forms that are usually used as nouns is higher than that of words usually used as verbs, regardless of whether the form in question is being pronounced as a noun or a verb. They conclude that the significant Dominance effects suggest that speakers maximize the difference between noun- and verb-dominant words in conformity with the lexical distribution of English in which the majority of bisyllabic nouns are stressed on the first syllable and the majority of bisyllabic verbs on the second syllable.

We investigate the neural substrates of bisyllabic noun-verb homophones in English using near-infrared spectroscopy (NIRS). We used 16 bisyllabic noun-verb ambiguous pairs in Sereno & Jongman (1995). Note the grammatical categories of the words are never ambiguous in the sentence contexts. They are similar in meaning. The stimuli were presented randomly to the 6 adult (both male and female) native English speakers. The subjects were instructed to comprehend the auditorily presented sentences silently. The changes in hemoglobin (Hb) concentrations and their oxygeneration levels in the frontal and temporal lobes in the left hemisphere and the right homologous areas were recorded using NIRS systems (ETG-4000; Hitachi, Tokyo, Japan). The probe pads were positioned onto the subjects' bilateral frontal and temporal



areas as shown in the left figure. Green squares indicate the channel positions. Red probes are emitters and blue ones are detectors. The numbers of the channels 1-22 correspond to the numbers 1-22 of the time course of hemoglobin responses in the left and right hemispheres in Figures 1-4.

Figure 1 shows the average values of the oxy-Hb changes of the 6 English subjects in the frontal and temporal lobes in the left hemisphere and the homologous areas in the right hemisphere for the forestressed nouns in solid lines and the forestressed verbs in broken lines in noun-dominant words. We find anatomical-behavioral correlations, with a left frontal cortical network activated for verbs, which is marked with a solid circle, and greater activation in the temporal regions for nouns, which is marked with a dotted circle. In the neuroimaging literature, there is a substantial corpus of studies asking whether nouns and verbs are represented in different neural substrates, or whether both categories are processed in the same anatomical area but with functional differences between them.

Our results show that noun and verb categories are represented in different neural substrates in the left hemisphere. We argue, following Shapiro & Caramazza (2003), that information about a word's grammatical category is represented independently of its meaning. Semantic information about nouns and verbs is stored in the frontal and temporal lobes, in which categorical distinctions are explicitly represented when nouns and verbs are activated: the greater activation of nouns than verbs in the temporal regions and the greater activation of verbs than nouns in the frontal regions. In the right hemisphere the activation of highly frequent nouns in noun-dominant pairs occurs in the frontal and temporal regions. As for the backstressed nouns and backstressed verbs in noun-dominant words, we find little activation of the verbs and the nouns in the frontal and temporal lobes both in the left and right hemispheres.

Table 1 shows the mean, sum of the mean, sum of the square of the mean of oxy-Hb (mM·mm) of the nouns and verbs during the period of 5s of a stimulus and 19s of the intervals between stimuli at the channels listed in the parentheses by the 6 English subjects in the frontal and temporal lobes of the left hemisphere for the forestressed nouns and verbs in noun-dominant pairs. The degree of freedom and the calculated value of *t* are given. A value of 2.92 for the frontal lobe and 1.717 for the temporal lobe are needed in a directional test for significance at the $p \le 0.05$ level. Since our calculated values are larger than this, we reject the null hypotheses, and conclude that the means of the nouns and verbs do differ both in frontal and temporal lobes.

Table 1. Significant levels in the frontal and temporal lobes for the forestressed noun-dominant pairs.

frontal area (13, 17)					temporal ar	ea (3, 12, 21	, 7, 16, 11, 20,	6, 15, 10,	19, 14)	
noun	0.00065	0.0013	0.0000335	df_2	+-2.07	0.0137916	0.1655	0.0026962	df_22	t=4.50
verb	0.01605	0.0321	0.0005362	ui–2	1-2.97	0.001475	0.0177	0.0005769	ui=22	t=4.39

Figure 2 shows the verb dominant backstressed nouns and verbs. Verbdominant forestressed nouns and verbs show similar patterns of activation. We find the activation of both forestressed and backstressed verbs in the frontal lobe of the left hemisphere, which is marked with a solid circle. Furthermore, we find the decrease to the negative value in the oxy-Hb for both forestressed and backstressed verbs in the frontal lobe, which is marked with a dotted circle. Task-induced deactivation occurs because certain types of neural processes active during passive states are interrupted when subjects are engaged in effortful tasks (Binder et al. 2009). In the right hemisphere the activation of highly frequent verbs in verb-dominant pairs occurs both in positive and negative values in the frontal regions.

In Table 2 we give the same calculations of the oxy-Hb as in Table 1 for both forestressed and backstressed nouns and verbs in the frontal lobe of the left hemisphere in verb-dominant pairs. A critical value of significance with degree of freedom of 2 at the $p \le 0.05$ level is 2.92, and we conclude that the means of the nouns and verbs do differ. We find the statistically greater activation of verbs than nouns in stressed syllables in the frontal lobes in verb-dominant pairs



Figure 1 The oxy-Hb changes in the range of 0.10 and -0.10 mM·mm at the 22 channels in noundominant forestressed nouns and verbs in English (The sentence is within 5s, and the interval between sentences is 19s.)



Front

Right hemisphere



Figure 2 Verb-dominant backstressed nouns and verbs in English

to maximize the difference between noun- and verb-dominance in perception, though the amplitudes in stressed syllables in verb-dominant pairs are lower than the amplitude of the first syllable of the forestressed noun-dominant pairs in production. Note that the value of t is only slightly higher than the critical value in the frontal lobe for the forestressed noun-dominant pairs (see Table 1). We suggest that the cortical representation of speech does not merely reflect the external acoustic environment, but instead gives rise to the perceptual aspects relevant for the listener's intended goal.

Table	2.	Significant	levels in	the	frontal	lobe	for the	verb-	dominant	pairs
1 aoie	<i>~</i> .	Significant	ievens in	une	montai	1000	ior the	1010	aomman	puno

frontal area (4, 8), forestressed						frontal are	ea (4, 8), backstr	ressed		
noun	-0.0124	-0.0248	0.000312	df_2	t=6.00	-0.0203	0.0406	0.0008771	df_2	+-2.25
verb	0.02765	0.0553	0.0016046	df=2	1-0.00	0.00062	0.0124	0.0001027	ui=2	1-5.55

3. Neural Substrates of Bisyllabic Homophones in Japanese

The 8 sentence pairs which contain bisyllabic intra-category noun-noun homophones (e.g. *kasa*'umbrella', 'bamboo hat') and verb-verb homophones (e.g. *waku*'boil', 'gush out') were presented auditorily to the 6 adult (both male and female) native Japanese speakers. In Japanese the majority of the homophones are composed of words from the same lexical categories.

Figure 3 shows the average values of the oxy-Hb changes of the 6 Japanese subjects in the frontal and temporal lobes in the left hemisphere for the first noun in solid lines and the second noun in broken lines of the homophonous noun pairs. We find the activation in middle temporal gyrus which is marked with a solid circle for the first noun, and the activation in inferior frontal gyrus and inferior temporal gyrus which are marked with a dotted circle for the second noun, because there is temporal ambiguity before the noun in Japanese SOV sentences. Figure 4 shows the homophonous verb pairs. The first verb and the second verb show similar patterns of activation. They are not ambiguous in the sentence contexts, because the noun is given before the verbs. Most importantly nouns and verbs are processed in the same anatomical area in the brain in Japanese.

4. Neural Substrates of Nouns and Verbs

Hurford (2014) discusses the evolution of nouns and verbs. Noun and verb are the basic major word classes, or syntactic categories. Some languages get close to making no distinction between nouns and verbs, while in other languages the distinction is clearly visible in the positions that words occupy in sentences. In this dimension some languages have more grammar than others. How did this most basic grammatical distinction evolve? Hurford suggests that the answer



Figure 3 Homophonous noun pairs in Japanese



Figure 4 Homophonous verb pairs in Japanese

lies in the central special function of communication in language: giving information about identified objects. In the most ordinary kind of simple sentence, the subject expression with a noun at its core, is the Topic, and the predicate expression, with a verb at its core, is the Comment. This is the original basis for the pervasive noun/verb distinction at the heart of the grammar of languages.

Kemenade & Los (2006) discuss, based on the syntactic and discourse properties of the adverbs *ba* and *bonne*, that Old English (OE) discourse organization was closely interwoven with syntactic organization. The transition of Middle English (ME) is marked by the elimination of the multiple topics in OE, and results in a more strictly syntactic organization of the clause.

We assume that the earliest human language had SOV word order (Newmeyer 2000, Goldin-Meadow et al 2008, Ogura & Wang 2012), and in OE the basic word order of SOV was intertwined with various types of word order including verb second motivated by discourse organization. Ogura & Wang (2012, 2014) assert that the evolution of the postnominal relative clause in English from independent sentences through the paratactic adjunctive stage to the embedded structures, and its product, the SVO order, brought about a strictly syntactic organization of the clause in ME.

With this transition, the periphrastic constructions of progressive, perfect, pluperfect and modal auxiliaries, definite article and periphrastic *do* arose due to the speakers' desire to be more specific and informative than was possible with

the older forms, and case and gender as concordial categories disappeared. Also we assert that nouns in the Topics and verbs in the Comments which were strongly triggered by the discourse considerations in OE were clearly specified in the positions that they occupy in sentences in ME. We suggest that nouns and verbs were processed in the same anatomical areas of frontal and temporal lobes before OE as shown in Japanese; after ME semantic information about nouns and verbs is stored in the frontal and temporal lobes, in which categorical distinctions are explicitly represented.

Japanese is SOV language, and nominal relations are expressed by postpositional particles. While the SOV order is basic, Japanese allows reordering of preverbal major constituents. Furthermore, the subject is not expressed in colloquial speech if it is known from the context. We, therefore, find that nouns and verbs are processed in the same anatomical areas of frontal and temporal lobes in Japanese.

5. Conclusion

We have investigated the evolution of homophones as a result of interaction between the speaker's production and the listener's perception. In so doing, we have demonstrated that noun and verb categories are represented in different neural substrates in Present-day English. Semantic information about nouns and verbs is stored in the frontal and temporal lobes, in which categorical distinctions are explicitly represented when nouns and verbs are activated. In contrast, both noun and verb categories are processed in the same anatomical area, the frontal and temporal lobes, in Japanese.

We have suggested that nouns and verbs are originally processed in the same anatomical area of the brain when the basic word order is SOV, and discourse organization is closely interwoven with syntactic organization as shown in OE and Japanese. When the SOV order is changed to SVO through embedding, a strictly syntactic organization of the clause appears. The noun/verb distinction is clearly visible in the positions that nouns and verbs occupy in sentences, and it is represented in different neural substrates in the brain as shown in Present-day English.

Acknowledgements

M. Ogura is supported by the grants from the Human Frontier Science Program and the Ministry of Education, Culture, Sports, Science and Technology of Japan. W. S-Y. Wang is supported by HKRGC-GRF number 14611615.

- Binder, Jeffrey, R. et al. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex* 19, 2767-2796.
- Ferrer i Cancho, R. and R. V. Solé (2003) Least effort and the origins of scaling in human language, *PNAS* 100, 788-791.
- Goldin-Meadow, S. et al. (2008). The natural order of events: how speakers of different languages represent events nonverbally". *PNAS*, 105, 9163-9168.
- Hurford, J. R. (2014). *The Origins of Language: A slim guide*. Oxford: Oxfod University Press.
- Kemenade, Ans van & Bettelou Los (2006). Discourse adverbs and clausal syntax in Old and Middle English. In A. v. Kemenade and B. Los (Eds.),*The Handbook of the History of* English (pp. 224-248). Oxford: Blackwell.
- Newmeyer, F. J. (2000). On the reconstruction of 'proto-world' word order. In C. Knight, M. Studdert-Kennedy and J. R. Hurford (Eds.), *The Evolutionary Emergence of Language* (pp.372-388). Cambridge: Cambridge University Press.
- Ogura, M. (forthcoming). Language Evolution as a Complex Adaptive System: A multidisciplinary approach to the history of English. New York: Oxford University Press.
- Ogura, M. & W. S-Y. Wang (2012). Ambiguity resolution and evolution of word order. In T. C. Scott-Phillips et al. (Eds.), *The Evolution of Language*, (pp. 274-281). Singapore: World Scientific.
- ----- (2014). Evolution of tense and aspect. In E. A. Cartmill et al. (Eds.), *The Evolution of Language*, (pp. 213-220). Singapore: World Scientific.
- Piantadosi, S. T. et al. (2012). The Communicative function of ambiguity in language. *Cognition* 122, 280-291.
- Sereno, J. A. & Jongman, A (1995). Acoustic correlates of grammatical class. *Language and Speech* 38(1), 57-76.
- Shapiro, K. & A. Caramazza (2003). The representation of grammatical categories in the brain. *Trends in Cognitive Sciences* 7(5), 201-206.
- Wang, W. S-Y. (2011). Ambiguity in language. Korea Journal of Chinese Language and Literature 1, 3-20.
- Zipf, K. (1949). Human Behavior and the Principle of Least Effort: An Introduction toHuman Ecology. Cambridge, MA: Addison-Wesley.

LANGUAGE AS A PARASITE TO EMOTIONAL SIGNALS

KAZUO OKANOYA

cokanoya@mail.ecc.u-tokyo.ac.jp Department of Life Sciences, The University of Tokyo, Tokyo, Japan

The evolution of speech poses a dilemma when viewed from the signaling theory in evolutionary biology (Searcy & Nowicki, 2005). The dilemma comes from dual nature of speech. First, the content of speech, linguistic expressions, is not honest. This is because the content is not always grounded with the reality and by compositional nature of words and sentences, expression is infinite. Furthermore, the cost of linguistic content is quite low (Lachmann et al, 2001). The speaker can easily transmit false information regarding with the cost/benefit of the hearer. Trust must be established between the speaker and hearer before taking the linguistic message but such process itself is costly (Dawkins & Guilford, 1991). Second, the act of speech including bodily movement, facial expression, prosody, loudness, fluency, and vocabulary are mostly honest signals for health, intelligence, and genetic fitness. The creativity of language rests on the "ungrounded-ness" of linguistic content that is open to false information. If so, why did such a dishonest signal evolved at all?

I will lay out a hypothesis on this issue. Acoustical communication in terrestrial animals perhaps started as noises contingent with breathing and jaw or respiratory gestures associated with predation or feeding. These noises and bodily movements were gradually ritualized, forming the "fixed action patterns" of motor actions reflecting intentional or emotional states (Newman, 2012). Especially, vocalizations associated with respiratory action became indicators of emotional states such as fear or contentment, or that of intentions such as attack or copulation. Animal calls were thus established as an honest indicator of the internal state.

In certain species, stochastic combinations of such calls were used by young animals to induce parental behavior, perhaps because they reminded parents of immature articulation. This effect was then utilized by male animals to attract, or to suppress escape behavior of females. Extremes of such vocalizations are songs, used for mate attraction and/or territorial defense by many species of birds and some species of whales and primates (Fitch, 2006).

Songs are an honest signal of vigor, since singing is costly in terms of nutrition, safety, and time. Songs came to be utilized also in non-sexual social contexts seen in, for example, gibbons (Clarke, Reichard, & Zuberb hler, 2006).

Sequences of song syllables and behavioral contexts were gradually associated through a mutual segmentation process (Merker & Okanoya, 2007; Okanoya & Merker, 2007) and proto-words emerged. Arbitrary combinations of proto-words referred to non-existing entities and thus the dishonesty of speech started. However, speech was always associated with honest signals such as emotional expression in voice or facial. It is also suggestive that although fundamental frequency can be modified rather easily, changing formant frequencies are not as easy in human speech, securing honesty of speech output (Pisanski et al, 2014).

Thus, the receiver could mostly judge the honesty of the speech content. In this way, language content was able to evolve as parasitic to emotional expressions associated with the act of speech.

Furthermore, the dishonesty of linguistic expression produced creativity, and eventually, cumulative and transmittable culture. These byproducts provided strong survival value to humans as species competing with larger predator animals. In this way, linguistic contents and speech expression evolved in humans because of the mutual dependency between dishonesty and creativity.

Acknowledgements

Supported in part by This study was supported by MEXT Innovative Areas #4903, JP17H06380 to K.O.

- Clarke, E., Reichard, U. H., & Zuberbühler, K. (2006). The syntax and meaning of wild gibbon songs. *PLoS ONE*, *1*(1), e73.
- Dawkins, M. S., & Guilford, T. (1991). The corruption of honest signalling. Animal Behaviour, 41(5), 865-873.
- Fitch, W. (2006). The biology and evolution of music: A comparative perspective. *Cognition*, 100 (1), 173-215.
- Lachmann, M., Szamado, S., & Bergstrom, C. T. (2001). Cost and conflict in animal signals and human language. Proceedings of the National Academy of Sciences, 98(23), 13189-13194.
- Merker, B., & Okanoya, K. (2007). The natural history of human language: Bridging the gaps without magic. *Emergence of communication and language*, 403-420.
- Newman, J. D. (2012). *The physiological control of mammalian vocalization*: Springer Science & Business Media.
- Okanoya, K., & Merker, B. (2007). Neural substrates for string-context mutual segmentation: a path to human language. *Emergence of communication and*

language, 421-434.

- Pisanski, K., Fraccaro, P. J., Tigue, C. C., O'Connor, J. J., Röder, S., Andrews, P. W., ... & Feinberg, D. R. (2014). Vocal indicators of body size in men and women: a meta-analysis. Animal Behaviour, 95, 89-99.
- Searcy, W. A., & Nowicki, S. (2005). *The evolution of animal communication: reliability and deception in signaling systems*: Princeton University Press.

SOCIO-ENVIRONMENTAL CONSTRAINTS IN CATEGORY LEARNING

KATARZYNA OKTABA*1,2 and DARIUSZ KALOCIŃSKI1

*Corresponding Author: katarzyna.oktaba@wp.pl
¹Institute of Philosophy, University of Warsaw, Warsaw, Poland
²Warsaw School of Economics, Warsaw, Poland

According to Christiansen and Chater (2008), language is primarily shaped by communicative and cognitive pressures. In response, Wallentin and Frith (2008) claim that these constraints might be insufficient for explaining, for example, why we find variability and sub-optimality in certain systems of linguistic categorization. As a possible remedy, they propose to look at social motivations in language learning, which may run orthogonally to pressures for optimality.

To investigate the plausibility of the above proposal, we modify the Lewis signalling game (Lewis, 1969), which provides a model of linguistic conventionalization as shaped by environmental constraints. To include social pressures in our model, the payoffs of the game depend also on the similarity between agents' behaviour, even if it runs contrary to the environmentally optimal strategy.¹

Our model consists of two players, M = 2 terms (signals) and N > M objects, thus giving rise to categories (Skyrms, 2010; Barrett, 2007). There is a stable partition of the objects into two classes (natural groupings). Each player associates objects with signals probabilistically. Objects that are associated with the same term comprise one of the player vague categories. Players engage in repeated exchanges with sender-receiver roles. Each round the sender has some object in mind, sends a corresponding signal and expects another example of the same class *according to his private categorization* in return (social coordination). In addition, if the initial object and the received one belong to the same class according to the external (natural) grouping, players receive a positive payoff (environmental pressure). Thus, the overall payoff is given by $r = w_S * s + w_N * n$, where s, n are payoffs for social and external alignment, respectively, and $w_S + w_N = 1$.

We investigate learning of categorizations for a 2-term/27-object game played by two communicating agents whose roles alternate within consecutive exchanges. We take Herrnstein reinforcement learning (HRL) (Herrnstein, 1970; Roth & Erev, 1995; Skyrms, 2008) and its smoothed version (SRL) (Barrett &

¹The model mentioned here originally appears in a BA thesis by the first author (Oktaba, 2017).

Zollman, 2009), with parameters $\delta = 0.9$ (high forgetfulness) and $\lambda = 25$ (high reactivity), as conceivable models of learning. The algorithms are modified appropriately to incorporate both environmental and social rewarding systems.

We test each algorithm in three conditions: *social* ($w_N = 0.1$), *environmental* ($w_N = 0.9$), and *mixed* ($w_N = 0.5$). Exemplary results are shown in Figure 1.



Figure 1. Mean success rates for HRL and SRL, social vs environmental condition, based on 100 simulations run across 2000 rounds. The mean success rate is calculated from the last 50 interactions.

After 2000 rounds, social condition yields the best success rates (HRL 95%, SRL 99%). Mixed condition results in lower values (HRL 81%, SRL 98%). The lowest success is obtained in environmental condition (HRL 72%, SRL 97%).

Prima facie, a plausible reading of the suggestions made by Wallentin and Frith (2008) is that the need to be identified as members of one group may run counter to environmental pressures. Our analysis does not support this interpretation. On the contrary, it turns out that higher social pressures support adaptation to a stable environment. This conclusion seems valid for both population- and, to a lesser extent, individual-based dynamics.² Moreover, SRL simulations give further evidence that forgetting improves learning (Barrett & Zollman, 2009).

Another reading of the Wallentin and Frith proposal, which we left largely open, is to consider groups that are socially biased towards themselves and indifferent to social pressures upon between-group interaction, or groups that are mutually antagonized. It seems likely that such groups will develop shared languages but not on the between-group level, particularly in view of similar research on dialect formation (Blythe, Jones, & Renton, 2016).

²HRL provides an approximation of two interacting groups: probability weights may be viewed as capturing relative numbers of agents in a population adhering to a given association. Individual-based dynamics is better captured by SRL, especially in view of high values of forgetfulness and reactivity, reflecting severe memory constraints and eager processing, respectively (Christiansen & Chater, 2016).

Acknowledgements

Dariusz Kalociński was funded by the Polish National Science Centre grant 2015/19/B/HS1/03292.

- Barrett, J., & Zollman, K. J. S. (2009). The role of forgetting in the evolution and learning of language. *Journal of Experimental & Theoretical Artificial Intelligence*, *21*(4), 293–309.
- Barrett, J. A. (2007). Dynamic partitioning and the conventionality of kinds. *Philosophy of Science*, 74(4), 527–546.
- Blythe, R. A., Jones, A. H., & Renton, J. (2016). Spontaneous Dialect Formation In A Population Of Locally Aligning Agents. In S. G. Roberts, C. Cuskley, L. McCrohon, L. Barcel-Coblijn, O. Fehr, & T. Verhoef (Eds.), *The Evolution of Language: Proceedings of the 11th International Conference* (*EVOLANGX11*). Online at http://evolang.org/neworleans/papers/19.html.
- Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, *31*(05), 489–509.
- Christiansen, M. H., & Chater, N. (2016). The Now-or-Never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*, 39, e62.
- Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior*, *13*(2), 243–266.
- Lewis, D. K. (1969). Convention. Cambridge, MA: Harvard University Press.
- Oktaba, K. (2017). Learning to Signal: Socio-environmental Constraints in Category Formation. Bachelor's thesis, Institute of Philosophy, University of Warsaw, Poland. (http://semanticsarchive.net/Archive/ TA10Tk0Y/Constraints_in_Category_Formation_Oktaba_ 18.pdf)
- Roth, A. E., & Erev, I. (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and economic behavior*, 8(1), 164–212.
- Skyrms, B. (2008). Signals. Philosophy of Science, 75(5), 489-500.
- Skyrms, B. (2010). Signals: Evolution, learning, and information. OUP Oxford.
- Wallentin, M., & Frith, C. D. (2008). Language is shaped for social interactions, as well as by the brain. *Behavioral and Brain Sciences*, 31(5), 536–537.

SIMPLICITY BIASES OPERATE AT ITEM AND SYSTEM LEVEL. AN ITERATED LEARNING STUDY

ALIKI PAPA¹, MIRIAM STADJE², SIMON KIRBY² AND MONICA TAMARIZ*¹

*Corresponding Author: M.Tamariz@hw.ac.uk ¹Psychology, Heriot-Watt University, Edinburgh, UK ²CLE, PPLS, The University of Edinburgh, Edinburgh, UK

Cognitive biases, amplified over repeated cultural transmission, leave their mark on the structure of behaviour including language (Kirby et al., 2007). A key learning bias is the preference for simplicity – specifically, compressible representations (Kirby et al. 2015; Regier et al., 2015). The emergence of simplicity over cultural transmission has been experimentally attested: on one hand, *individual items* such as drawings become less complex graphically and increasingly recognizeable under the effects of schemas, or attractors (Bartlett, 1932; Sperber, 1996; Tamariz & Kirby, 2015). On the other hand, *sets of items* become compressible when initially independent, idiosyncratic elements develop system-level properties such as compositionality in miniature artificial languages (Kirby et al., 2015) or systematicity across sets of melodies (Verhoef 2012), colour strings (Cornish et al. 2014), rhythmical structures (Ravignani et al. 2016), or visual patterns on a grid (Claidiere et al. 2014; Kempe et al. 2015).

This study systematically explores the simultaneous emergence of simplicity on multiple levels through cultural transmission in the graphical modality. Given the above results, we hypothesize, for repeatedly transmitted sets of drawings:

- H1. The emergence of system-level properties leading to between-chain divergence and within-chain convergence of the drawings over generations. We expect chain-specific system-level properties such as a recogniseable chain-specific style, which contribute to set-wise simplicity, to emerge.
- H2. The emergence of item-level properties. Compressibility will additionally be achieved at the item level through (a) progressive simplification of the

graphical forms and (b) Increased **recognizeability** of individual drawings driven by schemas, or attractors.

We ran 10 transmission chains of 20 generations, with each participant reproducing 10 initially abstract drawings. Overall increase in set-wise simplicity, measured through an odd-one out selection task by naive participants, was not statistically significant. However, visual inspection found clear unique system properties, or style, emerging in several of the chains (H1). At the item level, graphical forms became significantly more compressible over generations (our complexity measure was scanned file size in KB) (H2a), and recognizeable schemas such as letters and numbers emerged throughout (H2b).

Our results partially confirm the ubiquity of the simplicity bias operating in cultural transmission. They speak to theories that emphasise the role of learning biases (e.g. Kirby et al. 2015) or attraction (e.g. Sperber, 1996) in cultural evolution. Finally, they highlight that cognitive and learning biases can operate across cultural modalities (linguistic, musical, visual and graphical), supporting a domain-general view of the cultural evolutionary process.

References

Bartlett, F. C. (1932). Remembering. Cambridge: Cambridge University Press.

- Claidière, N., Smith, K., Kirby, S. & Fagot, J. (2014). Cultural evolution of systematically structured behaviour in a non-human primate. *Proc. R. Soc. B* 281, 20141541.
- Cornish, H., Smith, K & Kirby, S. (2013). Systems from sequences. Proceedings of the 35th Annual Cognitive Science Conference, 340-345.
- Kempe, V., Gauvrit, N. & Forsyth, D. (2015). Structure emerges faster during cultural transmission in children than in adults. *Cognition 136*, 247-254.
- Kirby, S., Dowman, M. & Griffiths, T. (2007). Innateness and culture in the evolution of language. PNAS, 104, 5241-45.
- Kirby, S., Tamariz, M., Cornish, H. & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition* 141, 87-102.
- Ravignani, A., Delgado, T. & Kirby, S. (2016) Musical evolution in the lab exhibits rhythmic universals. *Nature Human Behaviour, 1*:1.
- Regier, T., Kemp, C., Kay, P. (2015). Word meanings across languages support efficient communication. *Handbook of Language Emergence*, 87, 237-263.
- Sperber, D. (1996). Explaining culture. Oxford: Blackwell.
- Tamariz, M. & Kirby, S. (2014). Culture: Copying, compression, and conventionality. Cognitive Science, 39, 171-183.
- Verhoef, T. (2012). The origins of duality of patterning in artificial whistled languages. Language and Cognition, 4, 357-380.

A CONCEPTUAL ESCAPE FROM THE PERILS OF UNIVERSAL GRAMMAR

JOEL PARTHEMORE

Joel.parthemore@his.se Department of Cognitive Neuroscience and Philosophy University of Skövde, Skövde, Sweden

Noam Chomsky's notion of an innate universal grammar (see e.g. Hauser & Chomsky, 2002; Chomsky, 1976) is known to be problematic, not least for identifying recursion as the defining aspect underlying human language, given the extremely limited use of recursion in known languages and ongoing claims (see e.g. Evans & Levinson, 2009) that some languages lack recursive capacity altogether. Far more serious, perhaps, is the sense that universal grammar is delivered by *a priori* fiat and empirical evidence is all but completely lacking. As Evans and Levinson argue, few aspects of language, indeed, are, on closer examination, in any meaningful sense universal, which might seem to point against universal biological foundations. Although a couple decades ago Chomsky's views on language still dominated the field, recent years have seen such notions as universal grammar being increasingly called into question.

Whether taking the form of a universal grammar or not, innate linguistic capacity has another consequence that might seem undesirable: by suggesting that syntactically and grammatically structured language is uniquely human, it risks creating an absolute divide between human beings and other terrestrial species – a number of which, as the burgeoning field of comparative cognition has revealed over the past couple decades, show complex cognitive abilities, including abilities, such as spontaneous tool use (e.g., Smith *et al.*, 2012), planning for future deception (Osvath & Karvonen, 2012), and mirror self-recognition (Gallup, 1977; Prior *et al.*, 2008; Plotnik *et al.*, 2006; Reiss & Marino, 2000), that were previously thought strictly to be human. Some, including this author (Parthemore, 2013a, 2013b) have argued that commitment to an absolute divide is a relic of religious traditions, and that understanding what human beings have in common with other species is prerequisite to understanding how human beings *are* unique.

Unfortunately if one argues, as Daniel Dennett sometimes appears to (1995, 2008) and someone like Zoltan Torey (2009:46, 123) does quite explicitly, that

human-style language is necessary for having a mind (at least in anything remotely like the human sense), then either language is a prerequisite for mind (and for such characteristically "higher" mental capacities as cognitive agency) or the two must arise simultaneously. Either way, an innate language capacity of one form or another – presumably delivered by evolution – seems all but impossible to avoid. Language is not something that human beings worked out collectively; it is effectively hard-wired in. Such a conclusion will be hard to accept for anyone hewing more to the empiricist than the rationalist tradition.

Such philosophers as Albert Newen, Andreas Bartels (Newen & Bartels, 2007), and Colin Allen (1999) offer an alternative – one that resonates well with comparative cognition research. Conceptual agency¹ exists along a continuum where it is appropriate to attribute some substantial degree of it so soon as an agent, through its behavior, exhibits a capacity to categorize the world, create novel categories, apply those categories to new circumstances, and – most importantly, perhaps – express surprise when the resulting expectations get violated. None of that *prima facie* requires language. What language does (see e.g. Parthemore, 2014) is not to make conceptual agency possible in the first place but allow pre-existing concepts to be contemplated on much more abstract levels and new, abstract concepts to be created that were not previously possible - all of which allows the agent to step further back from the present moment and engage in longer and longer ranging mental time travel (MTT). That is to say that the Bischof-Köhler Hypothesis promoted by Thomas Suddendorf and Michael Corballis (e.g., 2007) is wrong - but only for the absolute divide it creates, not for its intuition that there is something unique about human MTT.

Of course if one views conceptual agency as existing along a continuum then one can, and should, argue about where to draw the line even as one should probably be prepared to shift that line depending on the present context of discussion: that is to say, the line should be pragmatic, not absolute. Viewing conceptual agency in this way is shown to fit in nicely with recent discussions from e.g. Merlin Donald (2001, 1993), Jordan Zlatev (2005; Zlatev et al, 2005) and others suggesting that mimesis (roughly, systematically structured imitative behavior) is one of the necessary steps, both ontogenetically and phylogenetically, on the road to language as a facility that arises naturally out of a cultural context, beginning from episodic memory (itself already conceptually structured), rather than anything innate. As Donald writes (1998:185): "humans are undoubtedly unique in their spontaneous invention of language and symbols; but, as I have argued elsewhere... our special advantage is more on the production side than on the conceptual side of the ledger. Animals know much more than they can express."

¹ Concepts may be understood here either as the building blocks of systematically and productively structured thought, or the abilities by which certain agents cognize in a systematic and productively structured fashion.

- Allen, C. (1999) Animal concepts revisited: The use of self-monitoring as an empirical approach. *Erkenntnis*, **51**(1): 33–40.
- Chomsky, N. (1976) On the nature of language. *Annals of the New York Academy of Sciences*, **280**: 46-57.
- Dennett, D.C. (1995) Animal consciousness and why it matters. *Social Research*, **62**: 691–710.
- Dennett, D.C. (2008) Kinds of Minds: The Origins of Consciousness. Basic Books.
- Donald, M. (2001). A Mind So Rare: The Evolution of Human Consciousness. London: W.W. Norton.
- Donald, M. (1998) Material culture and cognition: Concluding thoughts. In Renshaw, C. & Scarre, C. (eds.), Cognition and Material Culture: The Archaeology of Symbolic Storage. McDonald Institute for Archaeological Research.
- Donald, M. (1993) Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition. Harvard University Press.
- Evans, N. & Levinson, S.C. (2009) The myth of language universals: Language diversity and its importance for cognitive science, *Behavioral and Brail Sciences*, **32**(5): 429-448.
- Gallup, G.G., 1977. Self recognition in primates: a comparative approach to the bi-direction properties of consciousness. *American Psychologist*, **32**(5): 329– 338.
- Hauser, M.D., Chomsky, N. & Fitch, W.T. (2002) The faculty of language: What is it, who has it, and how did it evolve? *Science*, **298**(5598): 1569-1579.
- Newen, A. & Bartels, A. (2007) Animal minds and the possession of concepts. *Philosophical Psychology*, **20**(3): 283–308.
- Osvath, M. & Karvonen, E. (2012) Spontaneous innovation for future deception in a male chimpanzee. *PloS one*, **7**(5): e36782.
- Parthemore, J. (2014) The case for protoconcepts: Why concepts, language, and protolanguage all need protoconcepts. *Theoria et Historia Scientiarum*, 11: 159-178.
- Parthemore, J. (2013a) The 'final frontier' as metaphor for mind: Opportunities to re-conceptualize what it means to be human. In Dunér, D., Holmberg, G., Parthemore, J. & Persson, E. (eds.), *The History and Philosophy of Astrobiology: Perspectives on the Human Mind and Extraterrestrial Life* (67-92), Cambridge Scholars Publishing.
- Parthemore, J. (2013b) Reappraising the relationship between concepts and language: The view from philosophy of mind and cognitive science. Paper presented at SALC IV (Scandinavian Association for Language and Cognition), 13-15 June 2013, Joensuu, Finland.

- Plotnik, J., de Waal, F., Reiss, D. (2006). Self-recognition in an asian elephant. Proceedings of the National Academy of Sciences, 103(45): 17053–17057. <u>http://www.pnas.org/content/103/45/17053</u> (accessed 2018-01-26).
- Prior, H., Schwarz, A., Güntürkün, O. (2008) Mirror-induced behavior in the magpie (pica pica): evidence of self-recognition. *PLoS Biology*, 6(8): e202. <u>http://dx.doi.org/10.1371/journal.pbio.0060202</u>
- Reiss, D., Marino, L. (2000) Mirror self-recognition in the bottlenose dolphin: a case of cognitive convergence. *Proceedings of the National Academy of Sciences*, 98(10): 5937–5942.
- Smith, B.P., Appleby, R. & Litchfield, C.A. (2012) Spontaneous tool use: An observation of a dingo (*Canis dingo*) using a table to access an out-of-reach food reward. *Behavioural Processes*, 89(3): 219-224.
- Suddendorf, T. & Corballis, M.C. (2007) The evolution of foresight: What is mental time travel, and is it unique to humans? *Behavioral and Brain Sciences*, **30**(3): 299-313.
- Torey, Z. (2009 [1999]) The Crucible of Consciousness: An Integrated Theory of Mind and Brain. MIT Press.
- Zlatev, J., Persson, T. & Gärdenfors, P. (2005) Mimesis: The "missing link" in human cognitive evolution. Lund University Cognitive Studies. http://www.lucs.lu.se/LUCS/121/LUCS.121.pdf (accessed 2018-01-26).
- Zlatev, J. (2005) What's in a schema? Bodily mimesis and the grounding of language. In Hampe, B. & Grady, J.E. (eds.) From Perception to Meaning: Image Schemas in Cognitive Linguistics. Walter de Gruyter.
EVOLUTION OF PHONEME USAGE DRIVEN BY PHONETIC ROBUSTNESS FOR ERROR MINIMIZATION

INES PENA-NOVAS¹, MARCO ARCHETTI^{*1}

*Corresponding Author: m.archetti@uea.ac.uk ¹School of Biological Sciences, University of East Anglia, Norwich, U.K.

1. Introduction

Phonemes are, in principle, neutral with respect to meaning. Yet, it has long been known (at least since Dewey 1923) that certain phonemes are used more often than others by a 100-fold factor. What accounts for this phoneme usage bias?

Here we suggest that phoneme usage bias is driven in part by what we dub *phonetic robustness*, the capacity of a phoneme to reduce errors. While they are neutral with respect to meaning, phonemes have different probabilities of mutation (errors in articulation), and their mutants differ in their similarity with the original phoneme. Phonemes, therefore, differ in robustness — the capacity of reducing the probability of articulation errors and their effect on perception errors. These differences in robustness can lead to usage bias over time simply due to different probabilities of transmission. A similar effect has been studied in molecular evolution where synonymous codons (which are neutral at the protein level) are used with non-random frequencies because they differ in genetic robustness (Archetti 2004, 2006; Plotkin et al. 2004, 2006).

We propose a quantitative measure of phonetic robustness based on articulation and perception distances between phonemes; we show that phonetic robustness can lead to changes in phoneme usage over time in a deterministic theoretical model and in stochastic simulations; and we show that phonetic robustness can predict phoneme usage in English words.

2. Methods

Phonetic robustness. Robustness R_{p1} for phoneme p1 is the complementary value of the average of the perceptual distances $P_{p1,p2}$ from all phonemes p_2 weighted by the probability of mutation $(1-D_{p1,p2})$ to p_2

$$R_{p1} = 1 - \sum_{p2} P_{p1,p2} \left(1 - D_{p1,p2} \right)$$

where $P_{p1,p2}$ is a normalised measure of perception distance (in this study, the distance between the first two principal components of phonological similarity – Mielke 2012) and $D_{p1,p2}$ is a normalised measure of articulation distance (in this study, the distance between the first two principal components of vocal tract distance – Mielke 2012).

Phoneme usage. Phoneme abundance was taken from the British National Corpus (Leech et al. 2001). Phonetic translation and phoneme frequencies were calculated using the Carnegie Mellon University Pronouncing Dictionary transcribed into IPA.

Theoretical analysis. The equilibrium frequencies of all phonemes were found by calculating the leading eigenvector of the matrix $(1-D_{p1,p2})P_{p1,p2}/\sigma_{p1}$, where σ_{p1} is a normalizing factor corresponding to the sum of the frequencies (before normalization) of the mutants of phoneme *p*1. We analysed a model with learning, in which errors can be corrected, and one without learning (σ_{p1} =1). We also analysed the same model in simulations for stochastic populations.

3. Results

Phonetic robustness is correlated with phoneme usage. We found a significant correlation between phoneme usage and robustness (R=-0.62, p<0.001 for all words). The correlation changes only slightly with the part of speech. A negative correlation means that less robust phonemes are used more often.

Phonetic robustness can lead to the observed phoneme usage bias. A model with learning, in which errors can be corrected, leads to an increase in frequency of the least robust phonemes. Phoneme frequencies change over time and their equilibrium values are correlated with robustness (R=-0.65, p<0.001) and with phoneme usage observed in the BNC (R=-0.57, p<0.005).

4. Discussion

Our results suggest that, if larger mutations can be detected and corrected more easily whereas mild mutations can persist undetected and uncorrected, robust phonemes will decrease in frequency over time (as their mutants are more likely to be transmitted), whereas anti-robust phonemes will persist (because their mutants are corrected, reverting to the original) and therefore increase in frequency over time. These results are in line with analogous observations in evolutionary genetics, where anti-robust codons in protein-coding genes increase in frequency over time because their mutants are detected and corrected with a higher probability than the mutants of robust codons. Our results suggest that phonetic robustness can explain why phonemes are used with unequal frequencies is words, and therefore that phonetic robustness is a fundamental force driving the evolution of language.

- Archetti, M. (2004) Selection on codon usage for error minimization at the protein level. J. Mol. Evol. 59, 400–415.
- Archetti, M. (2006) Genetic robustness and selection at the protein level for synonymous codons. J. Evol. Biol. 19, 353-365.
- Dewey, G. (1923), *Relative frequency of English speech sounds*. Harvard University Press, Cambridge, MA.
- Leech, G., Rayson, P., Wilson, A. (2001) Word Frequencies in Written and Spoken English based on the British National Corpus. Longman, London, UK.
- Mielke, J., (2012) A phonetically based metric of sound similarity. *Lingua* 122, 145–163.
- Plotkin, J.B. Dushoff, J., Desai, M.M., Fraser, H.B., (2006) Codon usage and selection on proteins. J. Mol. Evol. 63, 553-635.
- Plotkin, J.B. Dushoff, J., Fraser, H.B. (2004) Detecting selection using a single genome sequence of *M. tuberculosis* and *P. falciparum. Nature* 428, 943– 945.

THE EVOLUTIONARY DYNAMICS OF LANGUAGE AS A FUNCTION OF DEMOGRAPHY

DIRK PIJPOPS^{1,2}, KATRIEN BEULS³, and FREEK VAN DE VELDE^{*2}

*Corresponding Author: freek.vandevelde@kuleuven.be
¹Research Foundation Flanders (FWO), Brussels, Belgium
²QLVL, University of Leuven, Leuven, Belgium
³AI Laboratory, Vrije Universiteit Brussel, Brussels, Belgium

1. Introduction

Lieberman et al. (2007) aimed to quantify the evolutionary dynamics of language by investigating the rise of the English regular past tense inflection, which they equated with the weak *-ed* suffix. Yet, their bold conclusion that "*the half-life of an irregular verb scales as the square root of its usage frequency: a verb that is 100 times less frequent regularizes 10 times as fast*" (Lieberman et al., 2007, p.713) has successively attracted criticism from scholars in the fields of historical and evolutionary linguistics. First, Carroll, Svare, & Salmons (2012) showed that this constant regularization rate does not hold true for the closely-related German language. Second, Cuskley et al. (2014) found that the rise of the English weak -ed suffix is not driven by forces endogenous to language, such as analogy, but rather by external forces, such as new verbs entering the language through language contact.

We will reassess the constant-rate controversy by (i) extending the methodological scope with agent-based modeling, and (ii) extending the number of languages going beyond the German-English distinction, adding Dutch.

Our results show that the constant rate does not hold. If language change is co-determined by external forces resulting in languages adapting to its niche (Lupyan & Dale 2016) this is exactly what one would expect, since English, Dutch and German have endured external pressures to a different degree. We will focus on the influence of demographic change. In particular, we investigate the growth of cities and the resulting koineization due to migration in the three

language areas since the Middle Ages. The three different degrees of urbanization have led to different degrees of dialect contact, which could in turn, as we will argue, lead to different regularization rates. To support this claim, we will present both empirical evidence from linguistic and demographic databases, as well as the results of a computational simulation.

2. Empirical data

2.1. Linguistic data

To obtain a clear picture of the linguistic situation, we included the data on English from Lieberman et al. (2007) and the data on German from Carroll et al. (2012), and complemented these with our own Dutch data. This enables us to track the development of the past tense system of these three languages over a 1000 year period (800-1800).

2.2. Demographic data

For the demographic data, we make use of the databases of Bairoch et al. (1988), De Vries (1984), and Mitchell (1998). In particular, we compare the population growth of the largest cities in the English, Dutch and German language areas in each particular time period from 800-1800. Historical research has shown that the exponential growth of urban population cannot be reduced to natural growth, but is driven by immigration as well, both of foreigners and of by a rural exodus from the larger agglomeration, leading to dialect contact. We then observed correlations between the success of the weak inflection and the amount of demographic upheaval.

3. Simulation

A correlation between a demographic and a linguistic trend does not automatically entail a causation between the former and the latter, however. To further substantiate our claim, we therefore turn to an agent-based computer simulation. In this simulation, agents store exemplars or tokens of what they hear (cf. Pijpops et al., 2015), rather than type states (cf. Colaiori et al., 2015), and use these to produce novel forms. We find that (i) the weak inflection does not require special status as the single regular inflection in order to explain the tendencies observed in reality; (ii) replacement of verbs can indeed cause a continued rise of the weak inflection, even after a stable equilibrium between weak and strong verbs has emerged, confirming Cuskley et al. (2014); and most importantly (iii) if our current understanding of language, as implemented in the simulation, is correct, demography does indeed affect the rise of the weak inflection.

Acknowledgements

We would like to thank the anonymous reviewers for their constructive feedback as well as our colleague Isabeau De Smet for her help in this study.

- Bairoch, P, Batou, J. & Chèvre, P. (1988). La population des villes européennes de 800-1850. Geneva: Librairie Droz.
- Carroll, R., Svare, R., & Salmons, J. (2012). Quantifying the evolutionary dynamics of German verbs. *Journal of Historical Linguistics*, 2(2), 153–172.
- Colaiori, F., Castellano, C., Cuskley, C., Loreto, V., Pugliese, M., & Tria, F. (2015). General three-state model with biased population replacement: Analytical solution and application to language dynamics. *Physical Review*. *E, Statistical, Nonlinear, and Soft Matter Physics*, *91*(1–1), 12808.
- Cuskley, C., Pugliese, M., Castellano, C., Colaiori, F., Loreto, V., & Tria, F. (2014). Internal and External Dynamics in Language: Evidence from Verb Regularity in a Historical Corpus of English. *Plos One*, 9(8), e102882.
- De Vries, J. 1984. *European Urbanization 1500-1800*. Cambridge, Mass.: Harvard University Press.
- Lieberman, E., Michel, J.-B., Jackson, J., Tang, T., & Nowak, M. (2007). Quantifying the evolutionary dynamics of language. *Nature*, 449(7163), 713–716.
- Lupyan, G. & Dale, R. 2016. Why are there different languages? The role of adaptations in linguistic diversity. *Trends in Cognitive Science*.
- Mitchell, B.R. 1998. International Historical Statistics: Europe 1750-1993. Springer.
- Pijpops, D., Beuls, K., & Van de Velde, F. (2015). The rise of the verbal weak inflection in Germanic. An agent-based model. *Computational Linguistics in the Netherlands Journal*, 5, 81-102.

LANGUAGE EVOLUTION AND THE EMERGENCE OF CONSTRUCTIONS IN INTERACTION

MICHAEL PLEYER

mpleyer@uni-koblenz.de

English Department, Universität Koblenz-Landau, Campus Koblenz, Koblenz, Germany

This talk explores the evolution of language from the perspective of dialogic construction grammar, dialogic syntax, emergent grammar and interactional usage-based approaches (Goldberg 2006; Du Bois 2014; Hopper 2015, Brône & Zima 2014; Pleyer 2017). These approaches all focus on "the emergence of grammatical patterning in on-line production, dialogically, and as a cooperative achievement" (Fried & Östman 2005: 1756). This talk argues that investigations of the evolutionary emergence of grammatical structure and constructions need to pay attention to this focus on the online, in-vivo interactional emergence of constructions and their subsequent conventionalization.

As I argue, this process of interactive constructionalization also presents a way how the first protoconstructions could have emerged in protolinguistic hominin communities.

In interaction, cognitive processes such as priming, conceptual pacts, intersubjective alignment, local schematization and routinization, microentrenchment and resonance activation can lead to the emergence of a temporary local network of constructions (Brône & Zima 2014). In the case of frequently recurring routines and practices, ad hoc constructions that are frequently microentrenched also acquire increasing degrees of entrenchment and conventionalization. Over time, this leads to these constructions emerging more easily and more frequently in interactions. This in turn subsequently leads to them becoming fully entrenched and acquire that status of grammaticalized and conventionalized constructions, or protoconstructions, proper. This is in line with the usage-based approach, which sees linguistic constructions as abstractions from frequently reoccurring instances of language – or in this case protolanguage - in interaction (Pleyer 2017).

It is therefore argued that the first conventionalized and community-wide protoconstructions are the result of frequently recurring processes of the interactive co-construction of online protoconstructional templates. These were then subject to conventionalization and entrenchment processes (Schmid 2017). The account presented here therefore assumes that ad hoc, temporary protoconstructios emerged first, and then subsequently and over time evolved into more conventionalized and entrenched protoconstructions. These in turn evolved into construction proper through processes of grammaticalization and constructionalization (Heine & Kuteva 2007; Traugott & Trousdale 2013). In this way, temporary, emergent communicative routines turned into an inventory of entrenched community-wide communicative routines - protolanguage -, which then evolved into a fully grammaticalized and conventionalized structured inventory of constructions shared by a community – language. Crucially, these early, temporary, emergent protoconstructional templates should not be seen as form-meaning pairings in the same way as human constructions. Instead, they are more like form-meaning potential pairings in which communicative routines still have a high-degree of semantic ambiguity and openness. In order to function as communicative units within an interaction, they are still heavily reliant on context. This semantic openness, or underspecification, would also characterize community-wide protoconstructional inventories. This view is consistent with analyses of the gesture systems of non-human primates, whose meaning and function is heavily context-dependent and underspecified (Liebal et al. 2014; Moore 2014). The gesture systems of non-human primates and the protoconstructional inventory possessed by early humans are therefore theorized to share many similarities. However, the human protoconstructicon would have been much more semantically and expressively powerful, as they were based on much more powerful capacities for metacognition, advanced sociocognitive capacities and perspective-taking (Scott-Phillips 2015, Tomasello 2008). In addition, early humans likely also possessed a "massive storage" capacity for a bigger protoconstructional inventory than is found in non-human primates (Hurford 2012; Pleyer & Lindner 2014).

They also likely possessed more sophisticated symbolization and schematization skills, which aided in the abstraction and generalization of protoconstructions into constructions proper (Penn et al. 2008; Deacon 1997; Thompson et al. 1997).

As this outline shows, then, dialogic and interactive approaches to the emergence of structure have much to offer for language evolution research (Pleyer 2017).

- Brône, G., & Zima, E. (2014). Towards a dialogic construction grammar: Ad hoc routines and resonance activation. *Cognitive Linguistics*, 25(3), 457-495. doi:10.1515/cog-2014-0027
- Deacon, T. W. (1997). *The Symbolic Species: The Co-Evolution of Language and the Brain*. New York: Norton.
- Du Bois, J. W. (2014). Towards a dialogic syntax. *Cognitive Linguistics*, 25(3), 359-410.
- Fried, M., & Östman, J.-O. (2005). Construction Grammar and spoken interaction: the case of pragmatic particles. *Journal of Pragmatics*, 37(11), 1752-1778.
- Goldberg, A. E. (2006). *Constructions at Work: The Nature of Generalization in Language*. Oxford ; New York: Oxford University Press.
- Heine, B., & Kuteva, T. (2007). *The Genesis of Grammar: A Reconstruction*. Oxford: Oxford University Press.
- Hopper, P. J. (2015). An Emergentist Approach to Grammar. In B. MacWhinney & W. O'Grady (Eds.), *The Handbook of Language Emergence* (pp. 314-327). Malden, MA: Wiley Blackwell.
- Hurford, J. R. (2012). *The Origins of Grammar: Language in the Light of Evolution II*. Oxford: Oxford University Press.
- Liebal, K., Waller, B. M., Burrows, A. M., & Slocombe, K. E. (2014). Primate Communication: A Multimodal Approach. Cambridge: Cambridge University Press.
- Moore, R. (2014). Ape Gestures: Interpreting Chimpanzee and Bonobo Minds. *Current Biology*, 24(14), R645-R647. doi:http://dx.doi.org/10.1016/j.cub.2014.05.072
- Penn, D. C., Keith J. Holyoak and Daniel J. Povinelli. (2008). Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences*, 31(2), 109-130.
- Pleyer, M. (2017). Protolanguage and mechanisms of meaning construal in interaction. *Language Sciences*, 63, 69-90. doi:https://doi.org/10.1016/j.langsci.2017.01.003
- Pleyer, M., & Lindner, N. (2014). Constructions, Construal and Cooperation in the Evolution of Language. In E. A. Cartmill, S. Roberts, H. Lyn, & H. Cornish (Eds.), *The Evolution of Language. Proceedings of the 10th Conference* (pp. 244-251). Singapore: World Scientific.
- Schmid, H.-J. ,ed. (2017). Entrenchment and the Psychology of Language Learning: How We Reorganize and Adapt Linguistic Knowledge.
 Washington, DC: American Psychological Association.
- Scott-Phillips, T. C. (2015). Speaking our Minds: Why Human Communication is Different, and How Language Evolved to Make it Special. Houndmills: Palgrave Macmillan.

Thompson, R. K. R., Oden, D. L., & Boysen, S. T. (1997). Language-naïve chimpanzees (Pan troglodytes) judge relations between relations in a conceptual matching-to-sample task. *Journal of Experimental Psychology: Animal Behavior Processes*, 23, 31-43.

Traugott, E. C., & Trousdale, G. (2013). *Constructionalization and Constructional Changes*. Oxford: Oxford University Press.

FROM WHERE TO WHAT: A NEUROANATOMICALLY BASED EVOLUTIONARY MODEL OF THE EMERGENCE OF SPEECH IN HUMANS

OREN POLIVA

polivaoren@hotmail.com

Psychology department, Bangor University, Bangor, United Kingdom

In the brain of primates, the auditory cortex connects with the frontal lobe via the temporal pole (auditory ventral stream; AVS) and via the inferior parietal lobe (auditory dorsal stream; ADS). The AVS is responsible for sound recognition, and the ADS for sound-localization, voice detection and integration of calls with faces. I propose that the primary role of the ADS in non-human primates is the detection and response to contact calls. These calls are exchanged between tribe members (e.g., mother-offspring) and are used for monitoring location. Detection of contact calls occurs by the ADS identifying a voice, localizing it, and verifying that the corresponding face is out of sight. Once a contact call is detected, the primate produces a contact call in return via descending connections from the frontal lobe to brainstem motor nuclei. Because the ADS of present day humans also performs speech production and repetition, I further propose an evolutionary course for the transition from contact call exchange to an early form of speech. In accordance with this model, structural changes to the ADS endowed early members of the genus Homo with partial vocal control. This development was beneficial as it enabled offspring to modify their contact calls with intonations for signaling high or low levels of distress to their mother. Eventually, individuals were capable of participating in yes-no question-answer conversations. In these conversations the offspring emitted a low-level distress call for inquiring about the safety of objects (e.g., food), and the mother responded with a high- or low-level distress call to signal approval or disapproval of the interaction. Gradually, the ADS and its connections with brainstem motor regions became more robust and vocal control became more volitional. Speech emerged once vocal control was sufficient for inventing novel calls.

1. Introduction

In the past five decades, gorillas, orangutans, chimpanzees and bonobos were shown capable of learning sign language (Blake, 2004; Gibson, 2011). An important cognitive distinction between the language used by humans and the language used by other apes is with the ability to ask questions. This was first noted by (Premack & Premack, 1984) who reported that, although their chimpanzee, Sarah, showed no difficulty answering questions or repeating questions before answering them, she never used the question signs for inquiring about her own environment. Jordania (2006), in his review of the literature, noted that other signing apes did not utilize questions and that their initiation of conversations was limited to commands (e.g., "me more eat") and observational statements (e.g., "bird there"). This absence of a

This paper is distributed under a Creative Commons CC-BY-ND license.

DOI:10.12775/3991-1.094

questioning mind is in direct contrast to human toddlers and children, who are renown for their incessant use of questions. My interpretation of this human-ape distinction is that during human evolution, we transitioned from the display of curiosity toward items that are present in our environment (i.e., observational statements) to curiosity toward items that are absent in our environment (i.e., WH questions). Developing curiosity about out of sight events and objects could thus explain the rapid migration of humans across the globe. Furthermore, this curiosity toward the unknown is the driving force behind scientific exploration and technological development. One could hence argue that it is the ability to ask that separates us from other animals and makes the human species unique.

Although no non-human primate has been reported to ask questions, they were reported to exchange calls for monitoring location (i.e., contact calls). For example, when a mother and her infant are physically separated, each emits in turn a call to signal the other their location. This emission of contact calls could therefore be interpreted as akin in meaning to the question "where are you?". If human communication and contact calls are related, it suggests that the preliminary urge to learn about the unknown is derived from infants and mothers seeking to reunite.



Figure 1. Dual stream connectivity between the auditory cortex and frontal lobe of monkeys and humans. The brain of the monkey (left) and human (right) is schematically depicted and displayed from the side. On the left panel of each species, the auditory cortex is schematically depicted on the supratemporal plane (orange frame in the right panel) and observed from above (with the parieto-frontal operculi removed). Blue colors mark regions affiliated with the ADS, and red colors mark regions affiliated with the AVS.

2. Two pathways of auditory processing

In the brains of primates, two pathways connect the auditory cortex with the frontal lobe (Figure 1): In the auditory ventral stream (AVS), the anterior superior temporal gyrus/sulcus connects with the middle temporal gyrus/temporal pole, which further connects with the inferior frontal gyrus (IFG). This pathway is well known for its

role in recognizing speech and sounds in both hemispheres (Dronkers et al., 2004; Hickok et al., 2007; Gow, 2012), and damage to its structures has been associated with the disorders auditory agnosia (Poliva et al., 2015) and semantic dementia (Noppeney et al., 2007). The auditory dorsal stream (ADS) connects the posterior auditory cortex (pSTG) with the Sylvian parietal-temporal junction (Spt) and inferior parietal lobe (IPL)/intra-parietal sulcus (IPS), which also connects with the IFG. As will be detailed below, in non-human primates this pathway is involved in sound localization, speaker identification and integration of calls with faces. In the left hemisphere of the human brain, this pathway is also responsible for speech production and repetition, perception and production of prosody and integration of spoken words with their facial movements. Herein, I argue that the ADS original role was for exchanging contact calls and that it was modification of this pathway that first enabled early members of the species homo (i.e., Hominans) to be able to ask questions, and later develop language.

3. The role of the ADS with the detection of contact calls

Direct evidence for involvement of the ADS in the detection and response to contact calls is demonstrated in a study that sacrificed marmoset monkeys and measured genomic expression of cFos protein. Activation was reported only in the ADS (areas CM-CL and IFG) while hearing and responding to contact calls (phee calls; Miller *et al.*, 2010). Monkeys sacrificed after only hearing contact calls or only emitting them showed neural activity in the same regions but to a much smaller degree (See also Simões *et al.*, 2010 for similar results in a study using the protein Egr-1).

The detection of contact calls likely occurs in 3 stages: 1. Identify the speaker 2. localize the direction of the call 3. Confirm that the face of the speaker is not present in that direction. Consistent with a role of the ADS in the detection of contact calls, the ADS has been implicated with all these functions. Recordings of neural activity from the auditory cortex of monkeys showed selectivity in the posterior, but not anterior auditory cortex, of sound localization (Benson *et al.*, 1981; Miller & Recanzone, 2009; Rauschecker *et al.*, 1995; Tian *et al.*, 2001; Woods *et al.*, 2006). Functional imaging studies in humans also correlated ADS' regions (IPL, IFG) with sound localization (Ahveninen *et al.*, 2006; Barrett & Hall, 2006; Brunetti *et al.*, 2005; De Santis *et al.*, 2007; Hart *et al.*, 2004; Krumbholz *et al.*, 2005; Warren & Griffiths, 2003). A role of the ADS in speaker identification was shown in an intra-cortical recording study of epileptic patients that associated activation in the pSTG with speaker discrimination (Lachaux *et al.*, 2007). Recording from the posterior insula (near pSTG) of monkeys also reported of activity during speaker

discrimination (Remedios *et al.*, 2009). Functional imaging of third trimester human fetuses demonstrated activation in area Spt when the hearing of voices was contrasted to pure tones, and of a sub-region of area Spt that is more selective to maternal voice than unfamiliar female voices (Jardri *et al.*, 2012). The role of the ADS in face-call integration is evident in an fMRI monkey study that reported activation only in the posterior auditory cortex during audio-visual stimuli integration (Kayser *et al.*, 2009). A study that recorded from neurons in A1 also showed that the activation is weaker while hearing contact calls than social calls (grunt call; Ghazanfar *et al.*, 2005). In humans, the neighboring pSTS is responsible integrating speech with lip-movements (McGurk effect; Beauchamp et al., 2010; Nath et al., 2010; Campbell, 2008).

In contrast to most calls, which are produced in limbic regions (Holstege et al., 2015), the IFG appears involved in the emission of contact calls. A study that recorded neural activity from the IFG of macaques reported of neural discharge prior to cued or spontaneous contact call production (coo calls), but not prior to production of vocalizations-like facial movements (i.e., silent vocalizations; Coudé *et al.*, 2011; see also Gemba *et al.*, 1999 for similar results). A tracing study in monkeys also reported of connections between IFG and brainstem motor nuclei (Jürgens & Alipour, 2002), which potentially could endow the animal with partial vocal control. Accordingly, behavioral studies of monkeys reported of ability to modify contact calls with intonations (Biben *et al.*, 1986; Sugiura, 1998) or of learning to use them in order to request objects (Hihara *et al.*, 2003). Anecdotal reports of more generalized volitional vocal control, albeit rudimentary, were also reported in apes (Hayes & Hayes, 1952; Hopkins *et al.*, 2007; Koda *et al.*, 2007; Koda *et al.*, 2015; Taglialatela *et al.*, 2003; Wich *et al.*, 2008).

4. The ADS and the emergence of speech

A role for the ADS in the emergence of human speech is implied in the fossil record. A study that reconstructed the endocranium of early *Hominins* noted that *Homo habilis*, but not any of its *Australopith* ancestors, is characterized by enlargement of the IPL and IFG, whereas the rest of the endocranium remains extremely similar to the endocranium of modern apes (Tobias, 1987). A diffusion tensor imaging study that compared the white matter of humans and chimpanzees also reported of significant strengthening in ADS connectivity, but not AVS connectivity (Rilling *et al.*, 2012). Evidence for a role of the ADS in the transition from mediating contact calls into mediating human speech also includes genetic studies that focused on

mutation to the protein SRPX2 and its regulator protein FOXP2 (Roll et al., 2010). In mice, blockage of SRPX2 or FOXP2 genes resulted in pups not emitting distress calls when separated from their mothers (Shu et al., 2005; Sia et al., 2013). In humans, however, individuals afflicted with a mutated SRPX2 or FOXP2 were reported with speech dyspraxia (Roll et al., 2006; Watkins et al., 2002). A PET imaging study of an individual with a mutated SRPX2 gene correlated this patient's disorder with abnormal activation (hyper-metabolism) along the ADS (pSTG-Spt-IPL; Roll et al., 2006). Similarly, an MRI study that scanned individuals with mutated FOXP2 reported increased grey matter density in the pSTG-Spt and reduced density in the IFG, thus further demonstrating abnormality in ADS' structures (Belton et al., 2003). A role for the ADS in mediating speech production in humans has also been demonstrated in studies that correlated a more severe variant of this disorder, apraxia of speech, with IPL and IFG lesions (Deutsch, 1984; Edmonds & Marquardt, 2004; Hillis et al., 2004; Josephs et al., 2006; Kimura & Watson, 1989; Square et al., 1997). The role of the ADS in speech production is also demonstrated via a series of studies that directly stimulated sub-cortical fibers during surgical operations (Duffau, 2008-review), and reported that interference in the left pSTG and IPL resulted in an increase in speech production errors, and interference in the left IFG resulted in speech arrest (see also Acheson et al., 2011; Stewart et al., 2001 for similar results using magnetic interference in healthy individuals). Damage to Spt-IPL was also correlated with impaired ability to repeat sentences and unfamiliar words (conduction aphasia; Selnes et al., 1985; Axer et al., 2001; Bartha and Benke, 2003; Dronkers et al., 2004; Baldo et al., 2008, 2012; Fridriksson et al., 2010; Buchsbaum et al., 2011). One study even reported that stimulation of the left IPL resulted with patients believing that they spoke, when they didn't, and IFG stimulation resulted with the patients unconsciously moving their lips (Desmurget et al., 2009).

Further support for the transition from contact call exchange to human speech are provided by studies of hemispheric lateralization (Petersen *et al.*, 1978). In one study, Japanese macaques and other old world monkeys were trained to discriminate contact calls of Japanese macaques, which were presented to the right or left ear. Although all the monkeys were capable of completing the task, only the Japanese macaques were noted with right ear advantage, thus indicating left hemispheric processing of contact calls. In a study replicating the same paradigm, Japanese macaques had an impaired ability to discriminate contact calls after suffering unilateral damage to the auditory cortex of the left, but not right, hemisphere (Heffner & Heffner, 1984). This leftward lateralization of contact call detection is

similar to the long established role of the human left hemisphere in the processing human language (Geschwind, 1965).

5. From vocal control to the first conversation

A possible route for the transition from contact call exchange to proto-speech was proposed by Dean Falk (2004). She argued that due to bipedal locomotion and the loss of hair in early Hominins, mothers were not capable of carrying their infants while foraging. As a result, the mothers maintained contact with their infant through a vocal exchange of calls that resembles contemporary "motherese" (the unique set of intonations that caregivers use when addressing infants). As previously suggested by another researcher (Masataka, 2009), such intermediate prosodic phase in the development of speech is consistent with evidence presented above that monkeys, to a limited extent, are capable of modifying their contact calls with intonations, and that apes are endowed with slightly more versatile vocal control. In the opening paragraph of this paper, I described the inability of apes to ask questions, and proposed that the ability to ask questions emerged from contact calls. Because the ability to modify calls with prosodic intonations likely further developed as we began conversing with questions, I expand Falk's and Masataka's views regarding the prosodic origins of vocal language, and propose that the transition from contact calls to prosodic intonations could have emerged as a means of enabling infants to express different levels of distress. In such a scenario, the modification of a call with intonations designed to express a high level of distress is akin in meaning to the sentence "mommy, come here now!". Hence, the modification of calls with intonations could have served as a precursor for the development of prosody in contemporary vocal commands. On the other hand, the use of intonations for expressing a low-level of distress is akin in meaning to the sentence "mommy, where are you?". Therefore, this use of prosody for asking the first question could have served as the precursor for pragmatically converting calls into questions by using prosody as well. This transition could be related to the ability of present-day infants of using intonations for changing the pragmatic utilization of a word from a statement to a command/demand ("MOMMY!") or a question ("mommy?"). This view is consistent with a longitudinal developmental study of toddlers, which reported of the toddlers utilizing prosodic intonations in their speech prior to construction of sentences (Snow, 1994). A study of speech perception in adults also demonstrated that our ability to discriminate questions sentences from statement sentences is dependent on analysis of prosodic intonations (Srinivasan et al., 2003). Evidence of the relationship between asking questions and processing in the ADS is demonstrated in a diffused tensor imaging and fMRI study (Sammler *et al.*, 2015), which reported the participation of both the ADS and AVS in the discrimination of mono-syllabic words into questions or statements. The researchers further showed that this discrimination was impaired while interference was induced with TMS in the pre-motor cortex of the ADS. Supporting the role of the ADS in the discrimination of questions and statements is the finding that patients with phonological dementia, who are known to suffer from degeneration along the ADS (Gorno-Tempini *et al.*, 2008; Rohrer *et al.*, 2010), were impaired in distinguishing whether a spoken word was a question or a statement (Rohrer *et al.*, 2012).

A possible route for the transition from emitting low-level distress calls to asking questions is by individuals starting to utilize the former to signal interest about objects in their environment. Given that both contact call exchange and contemporary speech are characterized with turn taking, early *Hominans* could have responded to the low-level distress calls with either high- or low-level distress calls. For example, when an infant expressed a low-level distress call prior to eating berries, his/her mother could have responded with a high-level distress call that indicated the food is dangerous or a low-level distress call that indicated the food is safe. Eventually, the infant emitted the question call and waited for an appropriate answer from their mother before proceeding with their intended action. This vocal exchange was most likely the precursor to yes-no question-answer conversations.

- Acheson, D. J., Hamidi, M., Binder, J. R., & Postle, B. R. (2011). A common neural substrate for language production and verbal working memory. *Journal of Cognitive Neuroscience* 23, 1358–1367. doi:10.1162/jocn.2010.21519.
- Ahveninen, J., Jaaskelainen, I. P., Raij, T., Bonmassar, G., Devore, S., Hämäläinen, M., et al. (2006). Task-modulated "what" and 'where' pathways in human auditory cortex. *Proc. Natl. Acad. Sci. U.S.A.* 103, 14608–14613.
- Axer, H., Keyserlingk, von, A. G., Berks, G., & Keyserlingk, von, D. G. (2001). Supra- and Infrasylvian Conduction Aphasia. *Brain and Language* 76, 317– 331. doi:10.1006/brln.2000.2425.
- Baldo, J. V., Katseff, S., & Dronkers, N. F. (2012). Brain regions underlying repetition and auditory-verbal short-term memory deficits in aphasia: Evidence from voxel-based lesion symptom mapping. *Aphasiology* 26, 338–354. doi: 10.1080/02687038.2011.602391.
- Baldo, J. V., Klostermann, E., & Dronkers, N. F. (2008). It's either a cook or a baker: Patients with conduction aphasia get the gist but lose the trace. *Brain and Language* 105, 134–140. doi:10.1016/j.bandl.2007.12.007.
- Barrett, D. J. K., & Hall, D. A. (2006). Response preferences for "what" and 'where'

in human non-primary auditory cortex. *NeuroImage* 32, 968–977. doi:10.1016/j.neuroimage.2006.03.050.

- Bartha, L., & Benke, T. (2003). Acute conduction aphasia: An analysis of 20 cases. *Brain and Language* 85, 93–108. doi:10.1016/S0093-934X(02)00502-3.
- Beauchamp, M. S., Nath, A. R., and Pasalar, S. (2010). fMRI-Guided Transcranial Magnetic Stimulation Reveals That the Superior Temporal Sulcus Is a Cortical Locus of the McGurk Effect. Journal of Neuroscience 30, 2414–2417. doi: 10.1523/JNEUROSCI.4865-09.2010.
- Belton, E., Salmond, C. H., Watkins, K. E., Vargha-Khadem, F., & Gadian, D. G. (2003). Bilateral brain abnormalities associated with dominantly inherited verbal and orofacial dyspraxia. *Hum. Brain Mapp.* 18, 194–200. doi:10.1002/ hbm.10093.
- Benson, D. A., Hienz, R. D., & Goldstein, M. H. (1981). Single-unit activity in the auditory cortex of monkeys actively localizing sound sources: spatial tuning and behavioral dependency. *Brain Res.* 219, 249–267.
- Biben, M., Symmes, D., & Masataka, N. (1986). Temporal and structural analysis of affiliative vocal exchanges in squirrel monkeys (Saimiri sciureus). *Behaviour*, 259–273.
- Blake, J. (2004). Gestural communication in the great apes. In The Evolution of Thought: Evolutionary Origins of Great Ape Intelligence. Cambridge University Press.
- Brunetti, M., Belardinelli, P., Caulo, M., Del Gratta, C., Penna, Della, S., Ferretti, A., et al. (2005). Human brain activation during passive listening to sounds from different locations: An fMRI and MEG study. *Hum. Brain Mapp.* 26, 251–261. doi:10.1002/hbm.20164.
- Buchsbaum, B. R., Baldo, J., Okada, K., Berman, K. F., Dronkers, N., D'Esposito, M., et al. (2011). Conduction aphasia, sensory-motor integration, and phonological short-term memory - An aggregate analysis of lesion and fMRI data. *Brain and Language* 119, 119–128. doi:10.1016/j.bandl.2010.12.001.
- Campbell, R. (2008). The processing of audio-visual speech: empirical and neural bases. Philosophical Transactions of the Royal Society B: Biological Sciences 363, 1001–1010. doi:10.1093/cercor/13.10.1034.
- Coudé, G., Ferrari, P. F., Rodà, F., Maranesi, M., Borelli, E., Veroni, V., et al. (2011). Neurons Controlling Voluntary Vocalization in the Macaque Ventral Premotor Cortex. *PLoS ONE* 6, e26822. doi:10.1371/journal.pone.0026822.s004.
- De Santis, L., Clarke, S., & Murray, M. M. (2006). Automatic and Intrinsic Auditory "What" and "Where" Processing in Humans Revealed by Electrical Neuroimaging. *Cerebral Cortex* 17, 9–17. doi:10.1093/cercor/bhj119.
- Desmurget, M., & Sirigu, A. (2009). A parietal-premotor network for movement intention and motor awareness. *Trends Cogn. Sci. (Regul. Ed.)* 13, 411–419. doi:10.1016/j.tics.2009.08.001.
- Deutsch, S. E. (1984). Prediction of site of lesion from speech apraxic error patterns. In *Apraxia of speech: Physiology, acoustics, linguistics, management.* 113–134.

- Dronkers, N. F., Wilkins, D. P., Van Valin, R. D., redfern, B. B., & Jaeger, J. J. (2004). Lesion analysis of the brain areas involved in language comprehension. *Cognition* 92, 145–177. doi:10.1016/j.cognition.2003.11.002.
- Duffau, H. (2008). The anatomo-functional connectivity of language revisited. *Neuropsychologia* 46, 927–934. doi:10.1016/j.neuropsychologia.2007.10.025.
- Edmonds, L., & Marquardt, T. (2004). Syllable use in apraxia of speech: Preliminary findings. *Aphasiology* 18, 1121–1134. doi:10.1080/02687030444000561.
- Falk, D. (2004). Prelinguistic evolution in early hominins: whence motherese? *Behav. Brain Sci.* 27, 491–503– discussion 503–83.
- Fridriksson, J., Kjartansson, O., Morgan, P. S., Hjaltason, H., Magnusdottir, S., Bonilha, L., et al. (2010). Impaired Speech Repetition and Left Parietal Lobe Damage. *Journal of Neuroscience* 30, 11057–11061. doi:10.1523/JNEUROSCI. 1120-10.2010.
- Gemba, H., Kyuhou, S., Matsuzaki, R., & Amino, Y. (1999). Cortical field potentials associated with audio-initiated vocalization in monkeys. *Neuroscience Letters* 272, 49–52.
- Geschwind, N. (1965). Disconnexion syndromes in animals and man. I. *Brain* 88, 237–294.
- Ghazanfar, A. A. (2005). Multisensory Integration of Dynamic Faces and Voices in Rhesus Monkey Auditory Cortex. *Journal of Neuroscience* 25, 5004–5012. doi: 10.1523/JNEUROSCI.0799-05.2005.
- Gibson, K. R. (2011). Language or protolanguage? A review of the ape language literature. In The Oxford Handbook of Language Evolution. Oxford University Press, USA.
- Gorno Tempini, M. L., Brambati, S. M., Ginex, V., Ogar, J., Dronkers, N. F., Marcone, A., et al. (2008). The logopenic/phonological variant of primary progressive aphasia. *Neurology* 71, 1227–1234. doi:10.1212/01.wnl. 0000320506.79811.da.
- Gow, D. W., Jr (2012). The cortical organization of lexical knowledge: A dual lexicon model of spoken language processing. *Brain and Language* 121, 273– 288. doi:10.1016/j.bandl.2012.03.005.
- Hart, H. C., Palmer, A. R., & Hall, D. A. (2004). Different areas of human nonprimary auditory cortex are activated by sounds with spatial and nonspatial properties. *Hum. Brain Mapp.* 21, 178–190. doi:10.1002/hbm.10156.
- Hayes, C., & Hayes, K. J. (1952). Imitation in a home-raised chimpanzee. J Comp Physiol Psychol 45, 450–459.
- Heffner, H. E., & Heffner, R. S. (1984). Temporal lobe lesions and perception of species-specific vocalizations by macaques. *Science* 226, 75–76.
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience* 8, 393–402. doi:10.1038/nrn2113.
- Hihara, S., Yamada, H., Iriki, A., & Okanoya, K. (2003). Spontaneous vocal differentiation of coo-calls for tools and food in Japanese monkeys. *Neuroscience Research* 45, 383–389. doi:10.1016/S0168-0102(03)00011-7.

- Hillis, A. E., Work, M., Barker, P. B., Jacobs, M. A., Breese, E. L., & Maurer, K. (2004). Re-examining the brain regions crucial for orchestrating speech articulation. *Brain* 127, 1479–1487. doi:10.1093/brain/awh172.
- Holstege, G., and Subramanian, H. H. (2015). Two different motor systems are needed to generate human speech. J. Comp. Neurol. 524, 1558–1577. doi: 10.1016/0304-3940(92)90692-Z.
- Hopkins, W. D., Taglialatela, J. P., & Leavens, D. A. (2007). Chimpanzees differentially produce novel vocalizations to capture the attention of a human. *Animal Behaviour* 73, 281–286. doi:10.1016/j.anbehav.2006.08.004.
- Jardri, R., Houfflin-Debarge, V., Delion, P., Pruvo, J.-P., Thomas, P., & Pins, D. (2012). Assessing fetal response to maternal speech using a noninvasive functional brain imaging technique. *International Journal of Developmental Neuroscience* 30, 159–161. doi:10.1016/j.ijdevneu.2011.11.002.
- Jordania, J. (2006). Who Asked the First Question? *The Origins of Human Choral Singing, Intelligence, Language and Speech*, p. 334–338.
- Josephs, K. A. (2006). Clinicopathological and imaging correlates of progressive aphasia and apraxia of speech. *Brain* 129, 1385–1398. doi:10.1093/brain/ awl078.
- Kayser, C., Petkov, C. I., & Logothetis, N. K. (2009). Multisensory interactions in primate auditory cortex: fMRI and electrophysiology. *Hear. Res.* 258, 80–88. doi:10.1016/j.heares.2009.02.011.
- Kimura, D., & Watson, N. (1989). The relation between oral movement control and speech. *Brain and Language* 37, 565–590.
- Koda, H., Nishimura, T., Tokuda, I. T., Oyakawa, C., Nihonmatsu, T., and Masataka, N. (2012). Soprano singing in gibbons. *Am. J. Phys. Anthropol.* 149, 347–355. doi:10.1002/ajpa.22124.
- Koda, H., Oyakawa, C., Kato, A., & Masataka, N. (2007). Experimental evidence for the volitional control of vocal production in an immature gibbon. *Behaviour*, 681–692.
- Krumbholz, K., Schönwiesner, M., Rübsamen, R., Zilles, K., Fink, G. R., and Cramon, von, D. Y. (2005). Hierarchical processing of sound location and motion in the human brainstem and planum temporale. *European Journal of Neuroscience* 21, 230–238. doi:10.1111/j.1460-9568.2004.03836.x.
- Lachaux, J.-P., Jerbi, K., Bertrand, O., Minotti, L., Hoffmann, D., Schoendorff, B., et al. (2007). A blueprint for real-time functional mapping via human intracranial recordings. *PLoS ONE* 2, e1094. doi:10.1371/journal.pone.0001094.
- Lameira, A. R., Hardus, M. E., Bartlett, A. M., Shumaker, R. W., Wich, S. A., & Menken, S. B. J. (2015). Speech-Like Rhythm in a Voiced and Voiceless Orangutan Call. *PLoS ONE* 10, e116136. doi:10.1371/journal.pone. 0116136.s003.
- Laporte, M. N. C., & Zuberbühler, K. (2010). Vocal greeting behaviour in wild chimpanzee females. *Animal Behaviour* 80, 467–473. doi:10.1016/j.anbehav. 2010.06.005.

- Masataka, N. (2009). The origins of language and the evolution of music: A comparative perspective. *Physics of Life Reviews* 6, 11–22. doi:10.1016/j.plrev. 2008.08.003.
- Miller, C. T., Mandel, K., & Wang, X. (2010). The communicative content of the common marmoset phee call during antiphonal calling. *Am. J. Primatol.* 72, 974–980. doi:10.1002/ajp.20854.
- Miller, L. M., & Recanzone, G. H. (2009). Populations of auditory cortical neurons can accurately encode acoustic space across stimulus intensity. *Proc. Natl. Acad. Sci. U.S.A.* 106, 5931–5935.
- Nath, A. R., and Beauchamp, M. S. (2012). A neural basis for interindividual differences in the McGurk effect, a multisensory speech illusion. NeuroImage 59, 781–787. doi:10.1016/j.neuroimage.2011.07.024.
- Noppeney, U., Josephs, O., Hocking, J., Price, C. J., & Friston, K. J. (2007). The Effect of Prior Visual Information on Recognition of Speech and Sounds. *Cerebral Cortex* 18, 598–609. doi:10.1093/cercor/bhm091.
- Perlman, M., & Clark, N. (2015). Learned vocal and breathing behavior in an enculturated gorilla. Anim Cogn 18, 1165-1179. doi:10.1007/ s10071-015-0889-6.
- Petersen, M. R., Beecher, M. D., Zoloth, S. R., Moody, D. B., & Stebbins, W. C. (1978). Neural lateralization of species-specific vocalizations by Japanese macaques (Macaca fuscata). *Science* 202, 324–327.
- Poliva, O., Bestelmeyer, P. E. G., Hall, M., Bultitude, J. H., Koller, K., & Rafal, R. D. (2015). Functional Mapping of the Human Auditory Cortex: fMRI Investigation of a Patient with Auditory Agnosia from Trauma to the Inferior Colliculus. *Cogn Behav Neurol* 28, 160–180. doi:10.1097/WNN. 000000000000072.
- Premack, D., & Premack, A. J. (1984). The Mind of an Ape. W. W. Norton.
- Rauschecker, J. P., Tian, B., & Hauser, M. (1995). Processing of complex sounds in the macaque nonprimary auditory cortex. *Science*.
- Remedios, R., Logothetis, N. K., & Kayser, C. (2009). An Auditory Region in the Primate Insular Cortex Responding Preferentially to Vocal Communication Sounds. *Journal of Neuroscience* 29, 1034–1045. doi:10.1523/JNEUROSCI. 4089-08.2009.
- Rilling, J. K., Glasser, M. F., Jbabdi, S., Andersson, J., and Preuss, T. M. (2011). Continuity, divergence, and the evolution of brain language pathways. Frontiers in evolutionary neuroscience 3. doi:10.3389/fnevo.2012.00011/abstract.
- Rohrer, J. D., Ridgway, G. R., Crutch, S. J., Hailstone, J., Goll, J. C., Clarkson, M. J., et al. (2010). Progressive logopenic/phonological aphasia: Erosion of the language network. *NeuroImage* 49, 984–993. doi:10.1016/j.neuroimage. 2009.08.002.
- Rohrer, J. D., Sauter, D., Scott, S., Rossor, M. N., & Warren, J. D. (2012). Receptive prosody in nonfluent primary progressive aphasias. *Cortex* 48, 308–316. doi: 10.1016/j.cortex.2010.09.004.

- Roll, P., Rudolf, G., Pereira, S., Royer, B., Scheffer, I. E., Massacrier, A., et al. (2006). SRPX2 mutations in disorders of language cortex and cognition. *Human Molecular Genetics* 15, 1195–1207. doi:10.1093/hmg/ddl035.
- Roll, P., Vernes, S. C., Bruneau, N., Cillario, J., Ponsole-Lenfant, M., Massacrier, A., et al. (2010). Molecular networks implicated in speech-related disorders: FOXP2 regulates the SRPX2/uPAR complex. *Human Molecular Genetics* 19, 4848–4860. doi:10.1093/hmg/ddq415.
- Sammler, D., Grosbras, M.-H., Anwander, A., Bestelmeyer, P. E. G., and Belin, P. (2015). Dorsal and Ventral Pathways for Prosody. CURBIO 25, 3079–3085. doi:10.1016/j.cub.2015.10.009.
- Selnes, O. A., Knopman, D. S., Niccumm, N., & Rubens, A. B. (1985). The critical role Wernicke's area in sentence repetition. *Ann. Neurol.* 17, 549–557.
- Shu, W., Cho, J. Y., Jiang, Y., Zhang, M., Weisz, D., Elder, G. A., et al. (2005). Altered ultrasonic vocalization in mice with a disruption in the Foxp2 gene. *Proc. Natl. Acad. Sci. U.S.A.* 102, 9643–9648. doi:10.1073/pnas.0503739102.
- Sia, G. M., Clem, R. L., & Huganir, R. L. (2013). The Human Language-Associated Gene SRPX2 Regulates Synapse Formation and Vocalization in Mice. *Science*. doi:10.1126/science.1245079.
- Simões, C. S. (2010). Activation of frontal neocortical areas by vocal production in marmosets. *Front.Integr.Neurosci.* 4. doi:10.3389/fnint.2010.00123.
- Snow, D. (1994). Phrase-final syllable lengthening and intonation in early child speech. J. Speech Hear Res.; 37(4): 831–840.
- Square, P. A., Roy, E. A., & Martin, R. E. (1997). Apraxia of speech: Another form of praxis disruption. In *Apraxia: The neuropsychology of action*. 173–206.
- Srinivasan, R.J., Massaro, D.W. (2003). Perceiving prosody from the face and voice: distinguishing statements from echoic questions in English. *Lang Speech*, 46(Pt 1): 1–22.
- Stewart, L., Walsh, V., Frith, U., & Rothwell, J. C. (2001). TMS Produces Two Dissociable Types of Speech Disruption. *NeuroImage* 13, 472–478. doi: 10.1006/nimg.2000.0701.
- Sugiura, H. (1998). Matching of acoustic features during the vocal exchange of coo calls by Japanese macaques. *Animal Behaviour* 55, 673–687.
- Taglialatela, J. P., Savage-Rumbaugh, S., & Baker, L. A. (2003). Vocal production by a language-competent Pan paniscus. *Int J Primatol* 24, 1–17.
- Tian, B., Reser, D., Durham, A., Kustov, A., & Rauschecker, J. P. (2001). Functional Specialization in Rhesus Monkey Auditory Cortex. *Science* 292, 290–293. doi: 10.1126/science.1058911.
- Tobias, P. V. (1987). The brain of Homo habilis: A new level of organization in cerebral evolution. Journal of Human Evolution 16, 741–761.
- Warren, J. D., & Griffiths, T. D. (2003). Distinct mechanisms for processing spatial sequences and pitch sequences in the human auditory brain. *Journal of Neuroscience* 23, 5799–5804.
- Wich, S. A., Swartz, K. B., Hardus, M. E., Lameira, A. R., Stromberg, E., &

Shumaker, R. W. (2008). A case of spontaneous acquisition of a human sound by an orangutan. *Primates* 50, 56–64. doi:10.1007/s10329-008-0117-y.

Woods, T. M., Lopez, S. E., Long, J. H., Rahman, J. E., & Recanzone, G. H. (2006). Effects of Stimulus Azimuth and Intensity on the Single-Neuron Activity in the Auditory Cortex of the Alert Macaque Monkey. J. Neurophysiol. 96, 3323– 3337. doi:10.1152/jn.00392.2006.

PINNIPEDS HAVE SOMETHING TO SAY ABOUT SPEECH AND RHYTHM

ANDREA RAVIGNANI^{*1,2,3}, MAXIME GARCIA⁴, STEPHANIE GROSS⁵, KOEN DE REUS¹, NIENKE HOEKSEMA¹, ANA RUBIO-GARCIA¹, and BART DE BOER²

 *Corresponding Author: andrea.ravignani@gmail.com
 ¹Research Department, Sealcentre Pieterburen, Pieterburen, Netherlands
 ²Artificial Intelligence Lab, Vrije Universiteit Brussel, Brussels, Belgium
 ³Language and Cognition Department, Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands
 ⁵ENES Lab/Neuro-PSI, CNRS UMR9197, University of Lyon/Saint Etienne, Saint-Etienne, France
 ⁵Institute for Terrestrial and Aquatic Wildlife Research, University of Veterinary Medicine Hannover, Foundation, Buesum, Germany

Understanding the origins and evolution of human speech requires multidisciplinary approaches. One approach consists in investigating animals' capacities for sound production. In particular, the function, ontogeny, mechanisms and flexibility of vocal production in other animals can shed light, by homology or analogy, on early stages of vocal production in the human lineage. Historically, comparative work has been performed on non-human primates or avian species. However, at least four other taxonomic groups show (i) greater phylogenetic proximity to humans than birds, and (ii) on average, more developed vocal production learning capacities than primates. These groups are pinnipeds (true seals, sea lions, fur seals, and walrus), bats, cetaceans, and elephants (Ralls et al., 1985; Janik & Slater, 2000; Reichmuth & Casey, 2014; Mathevon et al., 2017). Here we focus on pinnipeds and argue that, although this research lags decades behind avian and primate work, pinniped work has much to offer to understand the origins of human speech and music (Ravignani et al., 2016). In this contribution, we: (i) critically review available evidence on pinnipeds' capacities for vocal learning and rhythm (Reichmuth & Casey, 2014; Ravignani et al., 2016), (ii) present longitudinal

400

data on vocal development in harbour seal pups, including a case study of spontaneous vocal mimicry (de Reus, 2017), (iii) discuss preliminary evidence on pinnipeds' abilities for 'turn-taking' (Ravignani, in review), obtained from multitrack group recordings and playback experiments, and (iv) complement acoustic studies with anatomical findings on pinnipeds' larynx and vocal tract (Ravignani et al., 2017). We show how acoustic features in seal vocalizations change with age and sex, suggest that seals time their calls based on a 'selfish herd' mechanism (Hamilton, 1971; Ravignani 2014), and argue for anatomical similarities between the human and harbour seal phonatory apparati. Together, our findings suggest that pinnipeds' abilities to produce sounds, and time them precisely, are more developed than previously surmised (Mathevon et al., 2017).

Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant agreement No. 665501 with the research Foundation Flanders (FWO) (Pegasus2 Marie Curie fellowship 12N5517N awarded to A.R.), a visiting fellowship in Language Evolution from the Max Planck Society (awarded to A.R.), and ERC Grant [283435] ABACUS (awarded to B.d.B.). M.G. was supported by a postdoctoral fellowship from the Fyssen Foundation. We are grateful to M.Anichini, P.Cook, H.Honing, S.Kotz, M.Mendez-Arostegui, R.Sonnweber for advice and support.

- Hamilton, W. D. (1971). Geometry for the selfish herd. Journal of theoretical Biology, 31(2), 295-311.
- Janik, V. M., & Slater, P. J. (2000). The different roles of social learning in vocal communication. *Animal Behaviour*, 60(1), 1-11.
- Khan, C. B., Markowitz, H., McCowan, B. (2006). Vocal development in captive harbor seal pups, *Phoca vitulina richardii*: Age, sex, and individual differences. *The Journal of the Acoustical Society of America*, **120**(3):1684-1694.
- Mathevon, N., Casey, C., Reichmuth, C., & Charrier, I. (2017). Northern Elephant Seals Memorize the Rhythm and Timbre of Their Rivals' Voices. *Current Biology*.
- Ralls, K., Fiorelli, P., & Gish, S. (1985). Vocalizations and vocal mimicry in captive harbor seals, Phoca vitulina. *Canadian Journal of Zoology*, 63(5), 1050-1056.

- Ravignani, A., Fitch, W. T., Hanke, F. D., Heinrich, T., Hurgitsch, B., Kotz, S. A., Scharff, C., Stoeger, A. & de Boer, B. (2016). What pinnipeds have to say about human speech, music, and the evolution of rhythm. *Frontiers in neuroscience*, 10.
- Ravignani, A., Gross, S., Garcia, M., Rubio-Garcia, A., & de Boer, B. (2017). How small could a pup sound? The physical bases of signaling body size in harbor seals. *Current Zoology*.
- Ravignani, A. (2014). Chronometry for the chorusing herd: Hamilton's legacy on context-dependent acoustic signalling—a comment on Herbers (2013). *Biology letters*, 10(1), 20131018.
- Ravignani, A. (in review). Spontaneous rhythms in a harbor seal pup calls. *BMC Research Notes*.
- Reichmuth, C., & Casey, C. (2014). Vocal learning in seals, sea lions, and walruses. *Current opinion in neurobiology*, 28, 66-71.
- de Reus, K. (2017): Talking seals: Vocal development in Eastern Atlantic harbour seal pups (*Phoca vitulina vitulina*). *MSc.* Royal Veterinary College, University of London.

THE ROLE OF COMMUNITY SIZE IN THE EMERGENCE OF LINGUISTIC STRUCTURE

LIMOR RAVIV^{*1}, ANTJE MEYER^{,2}, and SHIRI LEV-ARI^{1,3}

*Corresponding Author: limor.raviv@mpi.nl
¹Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
²Radboud University, Nijmegen, The Netherlands
³Royal Holloway University of London, Egham, UK

1. Community Structure Affects Linguistic Structure

Over the last decade, several diachronic and typological analyses showed that the structure of languages spoken in exoteric societies is different from the structure of languages spoken in esoteric societies (e.g., Lupyan & Dale, 2010; Meir, Israel, Sandler, Padden & Aronoff, 2012; Nettle, 2012; Trudgill, 2009; Wray & Grace, 2007). These studies propose that different linguistic structures may emerge in different communities depending on their social properties. In particular, it has been argued that increased population size, sparser community structure and higher proportion of adult L2 learners in the community lead to morphological simplification. However, these three community properties are confounded in the real world, making it hard to evaluate their role separately. Additionally, the role of community size has not been experimentally tested.

2. The Current Studies

In the current studies, we focus on one property of community structure, namely, population size, and experimentally test the effect of different community sizes on the emergence of structure using a novel group communication game.

In Experiment 1 we tested six small communities of four participants and six larger communities of eight participants. Communities interacted in alternating dyads for seven rounds using an artificial language that was created by the participants on-the-go to refer to a meaning space that expanded with time. This was followed by a test round. At this point in time, all participants had the same amount of interaction and experience overall, but members of larger communities had less shared history with each other (i.e., each pair in the community has interacted fewer times). To equate the degree of shared history, larger communities were given seven additional communication rounds and an additional test round. In Experiment 2 we tested six more small communities that played as long as the larger communities in Experiment 1, and compared their performance to the larger communities across all 16 rounds to examine whether the differences found in Experiment 1 hold at a later point in time.

2.1. Results

We found that both small and larger communities developed compositional structure over time (measured as the average correlation between labels' string distances and meaning distances in the community, following Kirby, Tamariz, Cornish & Smith, 2015). Crucially, larger communities developed linguistic structure faster and more consistently than small groups. While there was no difference between the structure created by small and larger communities after eight rounds (seven communication rounds + test round), by the 16^{th} round, larger communities had more compositional structure than small communities (Figure 1). In addition, small and larger communities showed similar trends of stabilization and conventionalization by the 16^{th} round. Communicative success was not influenced by community size at any point in time. Finally, small communities showed significantly more variance than larger communities on all measures.

Together, and in line with previous typological studies (e.g. Lupyan & Dale, 2010), our findings demonstrate experimentally that population size can affect the formation of linguistic structure (as well as other linguistic properties), with larger communities developing structured languages faster and more consistently over time. These results highlight the role of the social environment in explaining patterns of linguistic diversity and trajectories of language change.



Figure 1. Linguistic structure by round number and community size. Only communities that played for 16 rounds are plotted (from both Experiment 1 & 2).

Acknowledgments

We wish to thank Caitlin Decuyper for programming the experiment, and Gary Lupyan, Mark Atkinson, Sean Roberts and Kevin Stadler for discussions and helpful input.

- Lupyan, G., & Dale, R. (2010). Language structure is partly determined by social structure. PloS one, 5(1), e8559.
- Meir, I., Israel, A., Sandler, W., Padden, C. A., & Aronoff, M. (2012). The influence of community on language structure: evidence from two young sign languages. Linguistic Variation, 12(2), 247-291.
- Wray, A., & Grace, G. W. (2007). The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. Lingua, 117(3), 543-578
- Trudgill, P. (2009). Sociolinguistic typology and complexification. In G. Sampson, D. Gil, and P. Trudgill (eds.), Language Complexity as an Evolving Variable. Oxford: Oxford University Press.
- Nettle, D. (2012). Social scale and structural complexity in human languages. Phil. Trans. R. Soc. B, 367(1597), 1829-1836.
- Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. Cognition, 141, 87-102.

CONCEPTUAL CATEGORIES SCAFFOLD VERBAL SEMANTIC STRUCTURE: A CROSS-CULTURAL STUDY OF CHILD HOMESIGN

LILIA RISSMAN*1, LAURA HORTON2, SUSAN GOLDIN-MEADOW2

*Corresponding Author: l.rissman@let.ru.nl

¹Center for Language Studies, Radboud University, Nijmegen, Netherlands ²Department of Psychology, University of Chicago, Chicago, IL, USA

In language evolution, formation of conceptual categories preceded formation of linguistic semantic categories (Hurford, 2007). This mapping from concepts to semantics is non-isomorphic, as particular languages categorize conceptual space in divergent ways (e.g. English *put in* is acceptable for both tight-fit and loose-fit relations, while Korean *kkita* encodes tight-fit relationships only; Choi & Bowerman, 1991). Despite this variation, are there crosslinguistic patterns in how words lexicalize conceptual space? We address this question analyzing how child homesigners from four different cultures describe instrumental events (e.g. cutting bread with a knife). Homesigners are congenitally deaf individuals who have not been taught a signed language. Despite growing up without structured linguistic input, these individuals use a gestural system ("homesign") to communicate (Goldin-Meadow, 2003). We find that homesign descriptions of instrumental events reflect categories present in adult English, Spanish and Mandarin, suggesting biases for how verbs encode the conceptual space of events, biases which may have been present over the course of language evolution.

English verbs such as *slice* and *write* encode the presence of an instrument, but *eat* and *open* do not (Koenig et al., 2003; Rissman et al., 2015; we label these <u>strong</u> and <u>weak</u> instrumental verbs, respectively). Rissman (2013) found that Spanish and Mandarin verbs fall into similar strong and weak instrumental categories as in English, suggesting that instruments are conceptually more salient in some events (e.g. slicing bread with a knife) than in others (e.g. eating pasta with a fork).

We tested this explanation by analyzing instrumental gestures in homesign. Nine homesigners from four cultures described cartoon pictures of instrumental events (1 from the United States: 3;5-4;10, 1 from Taiwan: 4;3-5;3, 1 from Nicaragua: 7;0-8;3, 6 from Guatemala: 8;6-11;4, 6;8, 10;10, 11;0-12;0, 6;11, 9;1-

406

9;10). For signs representing an action, we coded the morphosyntactic feature of *handshape type*: whether the sign had <u>handling</u> handshape (a grasping hand represents holding a knife) or <u>instrumental</u> handshape (a flat hand represents the shape of the knife); see Padden et al. (2013). We then asked adult native speakers of English, Spanish and Mandarin to describe the same cartoon pictures, and asked separate native speakers to categorize the verbs used by the first groups as either strong or weak, following Koenig et al. (2003) and Rissman (2013). Finally, we categorized each cartoon picture as to whether all three languages predominantly used strong instrumental verbs ("all strong"), as opposed to using predominantly weak instrumental verbs ("all weak"), or a mix of strong and weak instrumental verbs ("mix"). If English, Spanish, Mandarin and child homesign draw on similar instrumental event concepts, we predict that "all strong" pictures will be more likely to elicit <u>instrumental</u> handshape among the homesigners.

Figures 1 and 2 show the proportion of signs where a homesigner produced instrumental handshape, for each of the three picture types, for each homesigner. Our prediction was met: 8 out of 9 children were more likely to use instrumental handshape for "all strong" pictures. This suggests a basis for the strong/weak distinction that is not driven by language input. One possibility is that at a conceptual level, some events have more salient instruments than others, a conceptual categorization that may have influenced language evolution and led to common patterns of lexicalization across languages.



Figure 1. Guatemalan homesigners: proportion of signs with instrumental handshape, by whether the sign was describing an <u>all strong</u>, <u>all weak</u> or <u>mix</u> picture type. Total number of signs per child shown in panel label.



Figure 2. Nicaraguan, Taiwanese & U.S. homesigners (Panels 1, 2 & 3, respectively): proportion of signs with instrumental handshape, by whether the sign was describing an <u>all strong</u>, <u>all weak</u> or <u>mix</u> picture type. Total number of signs per child shown in panel label.

- Choi & Bowerman (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition.* 41(1), 83-121
- Goldin-Meadow (2003). *The resilience of language: What gesture creation in deaf children can tell us about how all children learn language*. New York, NY: Psychology Press.
- Hurford, J. R. (2007). The Origins of Meaning: Language in the Light of Evolution: OUP Oxford.
- Koenig, Mauner & Bienvenue (2003). Arguments for Adjuncts. *Cognition*. 89(2), 67-103.
- Padden, Meir, Hwang, Lepic, Seegers & Sampson (2013). Patterned iconicity in sign language lexicons. *Gesture*. 13(3), 287-308
- Rissman, Rawlins & Landau (2015). Using instruments to understand argument structure: Evidence for gradient representation. *Cognition.* 142(0), 266-290.
- Rissman, L. (2013). Event participant representations and the instrumental role: a cross-linguistic study. Johns Hopkins University, Baltimore, MD.

WORD FORM SHAPES ARE CULTURALLY SELECTED FOR INDICATING THEIR MORPHOLOGICAL STRUCTURE

Andreas Baumann, Christina Prömer, and Nikolaus Ritt*

*Corresponding Author: nikolaus.ritt@univie.ac.at Department of English and American Studies, University of Vienna, Austria

Listeners recognize morphologically complex word forms by their phonotactic shapes. For example, the final consonant clusters /md/ in seemed, /ld/ in filled, or /ks/ in socks function as clues that prompt decomposition. At the same time, experimental work (Post, Marslen-Wilson & Tyler 2008, cf. also Marslen-Wilson & Tyler 1997 & 1998) has shown that such recognition strategies are over-applied to word forms that are not actually complex, but simply happen to be shaped like complex ones. Thus, listeners attempt to decompose not only actual past tense forms such as *fill+ed* or *sign+ed*, or actual *plurals* such as *sock+s* but also simple words such as *build* or *find*, or *axe*, which delays processing, and leads listeners up garden paths. Obviously, however, such problems arise only when morphologically produced sound sequences have homophones among morphologically simple items. As we have seen, this is true of final clusters such as /nd/, /ld/ or /ks/. It is not true of clusters such as final /md/, /vd/, or /gz/, however. The latter occur only in past forms such as *seemed*, or *loved*, or in plurals, genitives, or 3sg forms such as eggs, Meg's or digs. Therefore, they signal complexity unambiguously and reliably.

If speakers are sensitive to the problems resulting from ambiguities between morphologically produced clusters (as in *fill+ed*) and lexically simple ones (as in *build*), they should be biased against the use of words that are ambiguous in that respect (Dressler & Dziubalska-Kołaczyk 2006; Dressler, Dziubalska-Kołaczyk & Pestal 2010). Since morphotactically ambiguous sequences are abundant in natural languages, however, the processing difficulties they produce do not seem to prevent successful communication and can only be slight. They nevertheless ought to be detectable in long-term language change, which results from vast numbers of communicative interactions and iterated learning events, known to

amplify even weak cognitive biases in language learning or language use (cf. Smith et al. 2017)

We have tested this hypothesis in a diachronic corpus study of highly ambiguous final clusters that arose through system-wide vowel deletion in unstressed syllables at the beginning of the Middle English period. In terms of data, it is based on the ECCE-database (ecce.univie.ac.at), which contains about 370.000 tokens of word forms ending in consonant clusters and covers the period from the middle of the 12^{th} to the middle of the 18^{th} century. The data are derived from the Penn Helsinki Parsed Corpora of Middle and Early Modern English (Kroch, Santorini & Delfs 2004, Kroch & Taylor 2000). In a statistical investigation of both type and token frequencies, we look for changes in the proportion of word forms that display either lexical or morphologically produced instances of final [sonorant]/d/, [sonorant]/t/, sonorant/z/, and sonorant/s/ clusters (as in *fill+ed* vs. *build*, *deal+t*,vs. *melt*, *name+*s vs. *Thames*, *mann+ys* vs. *prince*).

Our results show a significant trend towards cluster disambiguation: clusters in voiceless /t/ and /s/ have evolved to be increasingly indicative of morphological simplicity, while clusters ending in voiced /d/ and /z/ have come to signal complexity more reliably. The box chart in fig. 1 shows the difference for final sonorant+/t/ vs. sonorant-/d/ clusters on the level of types. The left box represents the distribution expected after schwa loss first produced final /nd/-clusters in words like *sinned* (/sind/ < /sinəd/) or *spilt* (/spilt/ < /spiləd/). The right one shows the distribution actually attested after the change.



Figure 1. Changes in the distribution of SON+/t/ and SON+/d/ across simple and complex forms.

As fig. 2 below shows difference between sonorant+/t/ and sonorant+/d/clusters came to correlate significantly more strongly with the difference between simple and complex word forms.



Figure 2. The evolving correlation between SON+/t/ vs. SON+/d/ and simple vs. complex words

One of the most interesting aspects of our results is that the disambiguation appears to be the cumulative effect of a set of rather diverse changes. They include sporadic sound changes (such as voicing of inflectional -s), morphological changes (the replacement of -Vnd participles by gerund forms ending in -ing), as well as the adoption and loss of specific lexical items. We interpret this to imply a selection bias for word forms that signal their morphological structure reliably. Such a bias may represent a plausible – and so far unrecognized -macro-condition on the cultural evolution of natural languages that constrains the actuation and the spread of changes on various levels of linguistic organization and gives direction to the long-term development of natural languages.

- Dressler, Wolfgang U. & Dziubalska-Kołaczyk, K. 2006. Proposing Morphonotactics. Wiener Linguistische Gazette 73. 69-87.
- Dressler, W. U., Dziubalska-Kołaczyk, K., & Pestal, L. 2010. Change and variation in morphonotactics. *Folia Linguistica Historica* 31. 51-68.
- Kroch, A., Santorini B., & Delfs L. 2004. Penn-Helsinki Parsed Corpus of Early Modern English. <u>http://www.ling.upenn.edu/hist-corpora/</u>.

- Kroch, A. & Taylor, A. 2000. Penn-Helsinki Parsed Corpus of Middle English. http://www.ling.upenn.edu/hist-corpora/.
- Marslen-Wilson, W. D., & Tyler, L. K. 1997. Dissociating types of mental computation. *Nature* 387. 592–594
- Marslen-Wilson, W. D., & Tyler, L. K. 1998. Rules, representations, and the English past tense. *Trends in Cognitive Science* 2, 428–435
- Post, B., Marslen-Wilson, W. D., & Tyler, L. K. 2008. The processing of English regular inflections: Phonological cues to morphological structure. *Cognition* 109. 1–17.
- Smith, K., Perfors, A., Fehér, O., Samara, A., Swoboda, K., & Wonnacott, E. 2017. Language learning, language use, and the evolution of linguistic variation. *Philosophical Transactions of the Royal Society B* 372, 20160051.
CHIELD: CAUSAL HYPOTHESES IN EVOLUTIONARY LINGUISTICS DATABASE

Seán G. Roberts

sean.roberts@bristol.ac.uk excd.lab, University of Bristol, Bristol, UK

Evolutionary linguistics is now a well established field with several conferences and its own journal. The ultimate goal of the field is to explain how complex communication systems emerge and change. A coherent, comprehensive explanation would involve a long chain of causal claims, stretching from genetics to cognition and from prehistorical adaptations to modern language change, supported by a range of methods from experiments to computer simulations. Because of the range of disciplines feeding into language evolution theories, producing such an explanation is a daunting task. In order to help this process, this paper presents a schema and implementation for a database of causal hypotheses about language evolution. Researchers can edit and contribute through a custom web application or through a GitHub repository.

1. Introduction

New databases and web technology are being used in many fields to synthesise knowledge. For example, the D-PLACE database (Kirby, Gray, Greenhill, Jordan, & al., 2016) integrates cultural, linguistic and phylogenetic data. Databases are also being created to collect hypotheses, too, such as the Explaining Human *Culture* database, a collection of over 3,000 hypotheses in cultural anthropology (Ember, 2016). Hypotheses are drawn from publications, and the database includes which variables were used, the statistical method and the main statistical results. The database is searchable by hypothesis or by variable, making it easy to find studies linking any two variables. Collections of studies like this can be used to guide research. For example, Metalab (Lewis et al., 2015) includes experimental results from 282 publications to support meta-analyses and power analyses in language acquisition paradigms. Collaboration tools are also helping to refine definitions and converge on hypotheses. For example, Glottolog (Hammarström, Forkel, & Haspelmath, 2017), a database of languages and language families, hosts its data on GitHub. Anyone can suggest edits and discuss issues in a simple web interface, allowing the research community to collaborate on maintaining and refining knowledge about linguistic history.

A similar resource for language evolution would be invaluable. The paper presents a schema and initial implementation for a database of causal hypotheses in evolutionary linguistics.

2. Motivation

The motivations for creating a database of causal hypotheses include:

Surveying the field. Language evolution is a very broad field, both in terms of scope and methods (Christiansen & Kirby, 2003), and surveying it is no easy task. Computational methods can help here (Bergmann & Dale, 2016), but the fundamental problem is simply the very large number of studies. Causal processes can be represented conveniently as graphical networks (Pearl, 2009), helping to visualise the field.

Converging on definitions. Part of the work of coding the database is to translate a hypothesis into an explicit series of causal links between variables. This forces transparent interpretations of theories and the use of common variable names. There will, of course, be disagreement on the interpretation of studies and on the terminology used for variables. However, if the debates can be centralised and directed towards concrete issues then this is a healthy process for a field.

Finding competing and supporting hypotheses. The database can identify competing explanations (alternative paths between variables or conflicting causal links). These are candidates for critical comparison studies. Similarly, the causal network could also identify evidence that supports a hypothesis, such as replications or tests using alternative methods. This aids a robustness approach to theory building (Irvine, Roberts, & Kirby, 2013).

Linking hypotheses together. The database could reveal some surprising links between theories, or identify missing or weakly supported links. It could also provide researchers with evidence for the preconditions for the topics they study, suggest wider downstream implications of their hypotheses or provide more detailed mechanisms that link higher-level concepts. Network analyses could identify 'broker' theories that bridge two areas. This would help extend theories and guide future research and collaboration.

Articulating causal processes. Even though causal arguments should be at the heart of any hypothesis investigation, coding articles for causal claims was often surprisingly difficult. Creating a visible framework for thinking about hypotheses as a network of causal processes will encourage more rigorous and transparent definitions of hypotheses. Using the schema below, it would be possible to publish a formal definition of the causal network alongside publications.

Research and teaching resource. The database will aid systematic literature review and provide an accessible entry point for students or researchers from outside of the field.

Given these motivations, there are several desiderata for a database of causal claims: it is openly accessible; the research community can contribute, edit and discuss issues; it should recognise contributors; causal claims can be represented visually and interactively; and the type of support for the claim should be coded; entries should be sourced widely and in an unbiased way.

3. Methodology

3.1. Framework

Causal claims can be represented as a directed graph (Pearl, 2009). Nodes represent variables and edges represent causal processes. The definition of variables is, at this point, left vague. This is because they might include a number of different kinds of concepts, depending on the research topic. For example, some variables might be concrete and measurable such as presence of a genetic allele, but others might represent higher-level concepts like a selection pressure for efficient communication. Also, variables might measure concepts on different scales, such as the age of an individual or the size of a population. While this is perhaps conceptually weak, in practice the interpretations are reasonably clear. Directed causal graphs can be easily visualised and analysed with a range of tools to discover weak, conflicting or supporting links (e.g. DAGitty, Textor, Hardt, & Knüppel, 2011).

3.2. Sources

The condition for entry into the database is that the hypothesis makes causal claims that relate to some part of the evolution of communication and that it is published in a peer-reviewed publication. Entry into the database does not mean that the hypothesis is correct nor widely accepted nor even empirically supported. The aim is not that the database be a single coherent, consistent theory of the evolution of communication, but a reflection of the field.

Existing digital databases will serve as initial sources of publications, such as the Language Evolution and Computation Bibliography (http://groups. lis.illinois.edu/amag/langev/), the Universals Archive (https:// typo.uni-konstanz.de/archive), the EvoLang conferences (http:// evolang.org/neworleans/) and relevant journals such as the Journal of Language Evolution and Interaction Studies. The research community can also contribute entries through a custom web application or directly through GitHub.

3.3. Coding scheme

An entry in the database encodes a single causal link between two variables. A minimal entry contains: bibtex reference for the source; label for variable 1; label for variable 2; type of causal relation; and the direction of the effect (positive or negative). A publication may be coded with multiple entries. The type of relation is drawn from table 1 (borrowing from the *lavaan* package in R, Rosseel, 2011).

The direction of the effect is necessary not only for interpreting the claim, but also so that causal claims from multiple studies can be integrated under the same variables (e.g. a process that increases morphological simplicity can be coded under a negative effect on morphological complexity).

Syntax	Meaning	Syntax	Meaning
$\overline{X > Y}$	A change in X causes a change in Y	X / > Y	X does not causally influence Y
$X \ll Y$	X and Y co-evolve	X >> Y	X is a necessary precondition for Y
$X \sim \sim Y$	X and Y are correlated	$X = \sim Y$	X is an indicator of (measured by) Y

Table 1. Causal relation syntax.



Figure 1. The current web interface for searching the database

Two studies might make claims about the same underlying concept, but measure it in different ways. To unify the theories, the causal link is represented as a latent variable: e.g. population size > morphological complexity (main link) morphological complexity = \sim presence of nominal case (indicator link) morphological complexity = \sim WALS feature score (indicator link)

Entries can be extended to include: Process: popular label for the process (e.g. "iterated learning"). Topic: e.g. phonetics, syntax etc. Stage: preadaptation, coevolution, cultural evolution, language change (Scott-Phillips & Kirby, 2010). Type: type of evidence (hypothesis (logical argument), review (other work), experiment, model, simulation). Subtype: subtype of evidence: e.g. iterated learning experiment, communication game etc. Confirmed: Whether the hypothesis was supported or not. Quote: A quote from the paper which states or clarifies the causal claim. Coder: Identity of the coder. These fields are important for the searchability of the database. For example, identifying the evolutionary stage at which the causal process applies helps to locate the link, but also to visualise the network of causal links.



Figure 2. A sub-graph from CHIELD linking population size and morphological complexity. Yellow: Ardell, Anderson & Winter (2016); Cyan, Purple: Atkinson, Smith & Kirby (2016); Pink: Bentz & Berdicevskis (2016); Red: Bentz & Winter (2013); Blue: Lupyan & Dale (2010); Green: Cuskley & Loreto (2016).

4. Current implementation

The database currently contains 222 causal links from 50 publications. The web interface can be accessed at http://chield.excd.org/, and the GitHub repository is live at https://GitHub.com/CHIELDOnline/CHIELD. The current interface (figure 1) allows users to interactively visualise different parts of the causal network, and submit their own links through a graphical interface. Coding of new links is guided by the interface's suggestions of variable labels already present in the database, helping to unify hypotheses. The data is hosted openly on GitHub, which also provides tools for curation, editing and debate.

Figure 2 shows part of the CHIELD database linking population size and morphological complexity (with an example of part of the coded data in table 2). While Bentz and Winter (2013) make a direct link between the proportion of adult learners and morphological complexity, two other studies discuss the intermediate step of the amount of the amount of linguistic input.

Table 2. An example of some entries in the database, summarising Lupyan & Dale (2010).

Variable 1	Relation	Variable 2	Cor	Туре	
population size	>	proportion of adult learners	pos	statistical	
proportion of adult learners	>	learning cost: morphology	pos	review	
learning cost: morphology	>	morphological complexity	neg	statistical	

5. Conclusion

This paper presented a schema and initial implementation of a database of causal hypotheses in evolutionary linguistics. Its aim is to provide an extendable resource for researchers. The major challenge is in the coding, both in terms of amount of time and coming to an agreement on interpretations and labels. The web interface tools and integration with GitHub are designed to address these challenges. However, there are also conceptual issues specific to language evolution (capturing arguments about the timing of the emergence of traits or properties such as population size having different connotations during preadaptation and language change). It is also unclear how observational work (e.g. animal communication) or arguments using evolutionary analogy fit in. Usefully visualising the network will also be a challenge, though there are many existing tools to help. Despite these difficulties, this paper argues that it is a worthwhile project which will has the potential for high impact in the field.

Acknowledgements

Supported by Leverhulme fellowship ECF-2016-435.

- Ardell, D., Anderson, N., & Winter, B. (2016). Noise in phonology affects encoding strategies in morphology. In S. Roberts, C. Cuskley, L. McCrohon, L. Barceló-Coblijn, O. Fehér, & T. Verhoef (Eds.), *The evolution of language: Proceedings of the 11th international conference (evolangx11)*.
- Atkinson, M., Smith, K., & Kirby, S. (2016). Adult language learning and the evolution of linguistic complexity. In S. Roberts & al. (Eds.), *The Evolution* of Language: Proceedings of the 11th Conference (EVOLANGX11).
- Bentz, C., & Berdicevskis, A. (2016). Learning pressures reduce morphological complexity: Linking corpus, computational and experimental evidence. In *Coling 2016* (pp. 222–232).
- Bentz, C., & Winter, B. (2013). Languages with more second language learners tend to lose nominal case. *Language Dynamics and Change*, *3*(1), 1–27.
- Bergmann, T., & Dale, R. (2016). A scientometric analysis of evolang: Intersections and authorships. In S. Roberts & al. (Eds.), *The evolution of language: Proceedings of the 11th Conference (EVOLANGX11).*
- Christiansen, M., & Kirby, S. (2003). Language evolution: The hardest problem in science? *Studies in the evolution of language*, *3*, 1–15.
- Cuskley, C., & Loreto, V. (2016). The emergence of rules and exceptions in a population of interacting agents. In S. Roberts & al. (Eds.), *The Evolution of Language: Proceedings of the 11th Conference (EVOLANGX11)*.
- Ember, C. R. (2016). *Explaining human culture*. New Haven, Ct.: Human Relations Area Files. http://hraf.yale.edu/ehc.

- Hammarström, H., Forkel, R., & Haspelmath, M. (2017). Glottolog 3.0. Jena: Max Planck Institute for the Science of Human History. http://glottolog.org.
- Irvine, L., Roberts, S. G., & Kirby, S. (2013). A robustness approach to theory building. In *Proceedings of the 35th CogSci* (pp. 2614–2619).
- Kirby, K. R., Gray, R. D., Greenhill, S. J., Jordan, F. M., & al. (2016). D-place: A global database of cultural, linguistic and environmental diversity. *PloS* one, 11(7), e0158391.
- Lewis, M., Braginsky, M., Bergmann, C., Tsuji, S., Cristia, A., & Frank, M. (2015). Metalab: A tool for power analysis and experimental planning in developmental research. In *Proceedings of 40th BUCLD*.
- Lupyan, G., & Dale, R. (2010). Language structure is partly determined by social structure. *PloS one*, *5*(1), e8559.
- Pearl, J. (2009). Causality. Cambridge university press.
- Rosseel, Y. (2011). *lavaan: an r package for structural equation modeling and more version 0.4-9 (beta).* Ghent University.
- Scott-Phillips, T. C., & Kirby, S. (2010). Language evolution in the laboratory. *Trends in cognitive sciences*, *14*(9), 411–417.
- Textor, J., Hardt, J., & Knüppel, S. (2011). Dagitty: a graphical tool for analyzing causal diagrams. *Epidemiology*, 22(5), 745.

INNOVATION, SELECTION AND THE EMERGENCE OF TRANSPARENT SIGNALS IN INTERACTION

Seán G. Roberts¹, Ashley Micklos², Justin Sulik³, and Hannah Little⁴

¹excd.lab, University of Bristol, UK, sean.roberts@bristol.ac.uk
 ²Linguistics Department, University of California, San Diego, USA
 ³Department of Psychology, Royal Holloway University of London, UK
 ⁴Department of Applied Sciences, University of the West of England, Bristol, UK

We review recent work in interactive experimental semiotics to discuss how biases in innovation and selection during interaction lead to the cultural evolution of transparent signals.

Recent studies suggest that individuals are not good at innovating transparent signals. For example, Sulik and Lupyan (2016) show that there are large individual differences in perspective taking abilities, with most participants in communication games being poor at choosing a signal that will be easy for their partners to interpret (though there are 'rare geniuses'). Verhoef, Roberts, and Dingemanse (2015) found that iconic signals could take generations to emerge, even with only 4 meanings and where the stimuli were designed to have obvious iconic mappings. Tamariz et al. (2017) found that innovations are equally likely to increase or decrease iconicity. While studies such as Tamariz et al. (2014) find that transparent signals are innovated in early generations, there are large individual differences in the ability to do this (though see Ortega, Schiefner, & Ozyurek, 2017; Schouwstra & de Swart, 2014). This would make innovation random (unbiased). This is supported to some extent by cross-linguistic studies arguing that iconicity in the lexicon both increases and decreases over cultural evolutionary time (Blasi et al., 2016). Even classic examples of individual innovation of transparent signals such as 'universal' sentence structures of creoles have recently been called into question (Blasi et al., 2017). How, then, do transparent linguistic conventions emerge?

One answer is that interaction provides the key mechanisms. Interaction can be thought of as an independent level of cultural evolution. In the broad model of genetic evolution (Dawkins, 1982; Hull, 1980), the gene is a *replicator* and an organism is a *vehicle* that interfaces with the environment to allow the replicator to replicate. According to Croft (2000) this model also applies to language: the word or phrase is a replicator and the individual speaker is the vehicle. However, this misses out a level between the individual's brain and the spoken phrases: turns at talk (sequences) in conversational interaction. This is highlighted in Buyn et al. (2016) which studies signers converging on a shared lexicon. They find that

frequency and transparency are good predictors of whether a form will survive, but that the type of sequence in which a form appears also matters. For example, a form produced in an explicit teaching sequence has a higher probability of surviving than a form produced in a simple statement. Several other studies also find that sequences such as repair can be loci of selection (Micklos, 2016) and can promote better communication in subsequent conversations (Mills & Redeker, 2016; Macuch Silva et al., 2017). This suggests that conversational sequences are also vehicles, at a different level to the individual speakers. The type of sequence promotes (or inhibits) the replication of the phrases within it, just as a particular type of cell within an organism influences the replication of the genes within it.

Indeed, Micklos (2017) suggests that specific interactional sequences such as other-initiated repair can transform the innovation process. If a partner does not understand, they are motivated to produce a new variant, which are often more transparent (biased innovation). That is, participants do not devote large processing effort to producing transparent signals until they encounter a communication problem, after which they do engage in more effortful perspective taking. Since people are good at *recognising* transparent signals (Sulik & Lupyan, 2016), once they arise they will be adopted and reproduced (Tamariz et al., 2014; Rogers & Fay, 2016).

A slightly different answer is that innovation is random, but that interaction simply provides inherent feedback about signal transparency. After negative feedback, participants innovate a new signal, which may be adopted if it is effective. Indeed, Tamariz et al. (2017) find that innovation is random but selection during interaction is biased towards adopting transparent signals.

The two possibilities could be compared experimentally or by looking at emergence dynamics. Random innovation and biased selection should exhibit punctuated equilibrium, while biased innovation should exhibit more gradual change. In either case, this would shield individuals from needing to apply perspective taking constantly, perhaps explaining the variation in perspective taking ability.

Acknowledgements

Thanks to Elizabeth Irvine for comments. SGR supported by Leverhulme fellowship ECF-2016-435.

- Blasi, D. E., Michaelis, S. M., & Haspelmath, M. (2017). Grammars are robustly transmitted even during the emergence of creole languages. *Nature Human Behaviour*.
- Blasi, D. E., Wichmann, S., Hammarström, H., Stadler, P. F., & Christiansen, M. H. (2016). Sound-meaning association biases evidenced across thousands of languages. *PNAS*, 201605782.

- Byun, K.-S., Roberts, S. G., De Vos, C., Levinson, S. C., & Zeshan, U. (2016). Content-biased and coordination-biased selection in the evolution of expressive forms in cross-signing. In *Proceedings of the 7th International Society for Gesture Studies Conference, Paris, France.*
- Croft, W. (2000). Explaining language change. Harlow, Essex: Longman.
- Dawkins, R. (1982). Replicators and vehicles. *Current problems in sociobiology*, 45, 64.
- Hull, D. L. (1980). Individuality and selection. Annual review of ecology and systematics, 11(1), 311–332.
- Macuch Silva, V., Roberts, S., Holler, J., & Özyürek, A. (2017). An experimental study of multimodal communication and early language emergence. Proceedings of Protolang 5.
- Micklos, A. (2016). Interaction for facilitating conventionalization: Negotiating the silent gesture communication of noun-verb pairs. In S. Roberts, C. Cuskley, L. McCrohon, L. Barceló-Coblijn, O. Fehér, & T. Verhoef (Eds.), *The evolution of language: Proceedings of the 11th international conference (evolangx11).*
- Micklos, A. (2017). Repair as a driver of innovation, improvement, and alignment in cumulative cultural evolution. In *Proceedings of the Cultural Evolution Society Conference, Jena.*
- Mills, G., & Redeker, G. (2016). Making invisible trouble visible: Self-repair increases abstraction of referring expressions. In *Cogsci 38*.
- Ortega, G., Schiefner, A., & Ozyurek, A. (2017). Speakers gestures predict the meaning and perception of iconicity in signs. In *the 39th annual conference* of the cognitive science society (cogsci 2017).
- Rogers, S. L., & Fay, N. (2016). Stick or switch: A selection heuristic predicts when people take the perspective of others or communicate egocentrically. *PloS one*, 11(7), e0159570.
- Schlag, K. H. (1998). Why imitate, and if so, how? *Journal of economic theory*, 78(1), 130–156.
- Schouwstra, M., & de Swart, H. (2014). The semantic origins of word order. *Cognition*, 131(3), 431–436.
- Sulik, J., & Lupyan, G. (2016). Failures of perspective taking in an open-ended signaling task. In S. Roberts & al. (Eds.), *The evolution of language: Proceedings of the 11th international conference (evolangx11).*
- Tamariz, M., Ellison, T. M., Barr, D. J., & Fay, N. (2014). Cultural selection drives the evolution of human communication systems. *Proceedings of the Royal Society of London B: Biological Sciences*, 281(1788), 20140488.
- Tamariz, M., Roberts, S. G., Martínez, J. I., & Santiago, J. (2017). The interactive origin of iconicity. *Cognitive Science*.
- Verhoef, T., Roberts, S., & Dingemanse, M. (2015). Emergence of systematic iconicity. In *Cogsci 37* (p. 2481-2486).

THE EMERGENCE OF FOCUS IN AN EXPERIMENTAL COMMUNICATION GAME

GARETH ROBERTS*1 and JON STEVENS2

*Corresponding Author: gareth.roberts@ling.upenn.edu
¹Department of Linguistics, University of Pennsylvania, Philadelphia, PA, USA
²Department of Linguistics, The Ohio State University, Columbus, OH, USA

We report an experiment investigating the emergence of focus, the prosodic or morphosyntactic marking of *critical elements* (Schmitz, 2008) in a sentence.

Stevens (2016) argued for a theory of focus based in information theory (Shannon & Weaver, 1949; Schmitz, 2008; Bergen & Goodman, 2015). Language users must deal with *noise* – the random deletion or alteration of parts of a signal. A solution is to compensate by adding redundancy (e.g., greater prosodic or morphosyntactic emphasis). However, redundancy costs both effort and time, so we should expect speakers to restrict redundancy to critical elements, particularly when effort and time pressures are high. (Redundancy on critical elements will be referred to here as *critical redundancy*, as compared with *non-critical* redundancy on other elements.) These factors should be expected to operate over multiple timescales. In a single interaction, speakers respond dynamically to perceived noise, time and effort pressures (Krauss & Weinheimer, 1964; Clark, 1996; Brennan & Clark, 1996). Developmentally, language learners acquire strategies for adding redundancy (Romaine, 1984). Over generations, such strategies can be expected to become grammaticalized as focus systems (Tamariz & Kirby, 2016).

We thus expect focus-like behavior to emerge and evolve in any communication system that involves sending messages under similar constraints and make the following predictions: (1) Overall message length should vary according to time and effort costs; (2) longer messages should differ from shorter messages not only with respect to length – shorter messages should also have lower proportions of non-critical redundancy; (3) critical redundancy should be higher when noise is higher, both in an absolute sense and in a relative sense (more critical than noncritical redundancy); (4) unless noise and time pressures actually prevent accurate communication, communicative accuracy should remain relatively constant, because focus is designed to help maintain accuracy under different conditions.

We tested these predictions experimentally by having participants play a simple communication game. Players sat separately and took turns to be "Sender" or "Receiver". The Sender would see three grids, two with line figures (Figure 1; in



Figure 1. Sender's screen.

half the trials, the line figures overlapped by five cells), one selected in green. The third grid was blank. The Sender's task was to communicate the selected grid to the Receiver by clicking on as many (or as few) cells as they liked in the blank grid. The Receiver was then shown the same three grids and had to choose the correct line figure. Both participants were told if the Receiver chose correctly.

There were six conditions based on manipulating effort, noise, and time constraints (Table 1). For a cell to be sent, a Sender had to click it 15 times in the *High effort condition* and five in the *Low effort conditions*. In the Noise conditions, any clicked cell would be sent with a probability of $1 - (1 - d)^n$, where *n* equals the number of clicks made on the cell. Two values were used for *d*: 0.1 in the *High noise condition* and 0.4 in the *Low noise conditions*. (There were no 5s High effort or High noise conditions, as the time limit would restrict the number of clicks too much.)

Table 1. Experimental conditions

Time limit	Effort	Noise
30 seconds	High effort	High noise
30 seconds	Low effort	Low noise
5 seconds	Low effort (5s)	Low noise (5s)

Results and discussion. All predictions were supported. Message length (i.e., the number of cells sent) was greater than necessary in all conditions and varied according to noise and effort levels, remaining constant in the Low effort and Low noise conditions, and declining over time in the High effort and noise conditions. Per-cell click rate gave a measure of emphasis added to different line segments. The proportion of clicks devoted to non-critical (as opposed to critical) redundancy correlated positively with overall click rate (r = 0.29, p < 0.001), but the correlation was stronger (r = 0.56) in the High effort and High noise conditions, where there was greater pressure on participants. The distribution of effort took noise into account, with critical redundancy higher when noise was higher ($\beta = 58.57, SE = 7.56, t = 7.75, p < 0.001$). Overall mean accuracy was 97%, and did not differ significantly between conditions, with one exception: It was lower in the Low noise (5s) condition ($\beta = -0.07, SE = 0.02, t = -2.81, p < 0.01$), likely due to participants' underestimating noise.

Our results lend support to the information-theoretic account and suggest how focus systems in language might arise. This experiment focused on strategies emerging in one generation; it is likely that these become grammaticalized through transmission between generations. In further experimental work, we plan to investigate this using an iterated learning approach (Kirby, Griffiths, & Smith, 2014).

- Bergen, L., & Goodman, N. D. (2015). The strategic use of noise in pragmatic reasoning. *Topics in Cognitive Science*, 7(2), 336–350.
- Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 1482–1493.
- Clark, H. H. (1996). Using language. Cambridge: Cambridge University Press.
- Kirby, S., Griffiths, T., & Smith, K. (2014). Iterated learning and the evolution of language. *Current Opinion in Neurobiology*, 28, 108–114.
- Krauss, R. M., & Weinheimer, S. (1964). Changes in reference phrases as a function of frequency of usage in social interaction—a preliminary study. *Psychonomic Science*, *1*(5), 113–114.
- Romaine, S. (1984). *The language of children and adolescents: The acquisition of communicative competence*. Oxford: Basil Blackwell.
- Schmitz, H.-C. (2008). *Accentuation and interpretation*. New York: Palgrave MacMillan.
- Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication*. Champaign, IL: University of Illinois Press.
- Stevens, J. (2016). Focus games. Linguistics and Philosophy, 39(5), 395-441.
- Tamariz, M., & Kirby, S. (2016). The cultural evolution of language. Current Opinion in Psychology, 8, 37–43.

COMPRESSIBLE KINSHIP TERMINOLOGIES ARE MORE LEARNABLE THAN LESS COMPRESSIBLE ALTERNATIVES

SARA ROLANDO¹, SIMON KIRBY¹, and KENNY SMITH^{*1} *Corresponding Author: kenny.smith@ed.ac.uk ¹Centre for Language Evolution, University of Edinburgh, UK

Different languages partition meanings into different semantic categories, labelled with words or morphemes. The scope of variation in these partitions is wide, as systems of semantic categories can differ in both the number of labels used and in the strategies used to group meanings into one category. However, this variation is constrained – not all theoretically-possible partitions are found in natural languages, and similar meanings are encountered in unrelated languages.

This pattern of constrained variation (Regier et al., 2015) has been attested in several domains, such as colour (Berlin and Kay, 1969), spatial terms (Ketharpal et al., 2009), numbers (Greenberg, 1978), and kin classification systems (Murdock, 1970), and clearly stands in need of explanation. Kemp and Regier (2012) propose that this constrained variation is a consequence of pressures for efficient communication (see also Regier et al., 2007; Xu and Regier, 2014; Carstensen et al., 2015). According to this view, category systems are shaped by two competing forces – the need for simplicity, as a learnable or efficient category system minimises cognitive load, and the need for accurate communication, as an efficient category system allows listeners to reliably reconstruct the meanings intended by the speaker.

Suggestive evidence supporting this hypothesis is provided by Kemp and Regier's (2012), who shows that kin categorization systems in natural languages exhibit a near-optimal trade-off between informativeness (ability to uniquely specify individuals in a family tree, weighted by the need probability of being required to refer to each individual) and simplicity (quantified by the length of the grammar underlying the kinship system). Here, we provide a mechanistic test of one of the key assumptions in Kemp and Regier's argument: we focus on

the pressure arising from language learning, and investigate whether simpler, more compressible kin categorization systems are indeed more learnable than less compressible alternatives.

To test this hypothesis, we ran an artificial language learning experiment in which participants were divided into two conditions and asked to learn two kinship terminologies that differed in their complexity but were matched in their expressivity and their similarity to English. Participants were familiarised with the family tree by introducing 3 members at a time and stating their relationship to each other by only using the primitives "child" and "parent". Once they had learned the set of family relationships to criterion, they were trained and tested on either the Simple or Complex kinship system (see Figure 1). Both kinship systems used 12 labels to individuate 16 possible relationships between relatives but differed in their complexity, where complexity was quantified using the metric proposed by Kemp and Regier (2012), specifically counting the number of rules necessary to describe the system (i.e. the length of the grammar).

As predicted under accounts which assume that simpler kinship systems are easier to learn, we found that participants in the Simple condition learned significantly faster than participants in the Complex condition. We also found that participants in both conditions tended to produce final languages with fewer labels than their input language, collapsing some distinctions encoded in their input, as predicted under accounts of learning where communicative functionality is sacrificed in favour of simplicity (e.g. Kirby et al., 2015).

These results support the claim that the learnability of a kin categorisation system (and by extension, other linguistic systems) depends on the complexity of the mental representation it requires, and that more compressible systems are more learnable than less compressible alternatives. In ongoing work we are testing whether the trade-off between simplicity and informativeness seen in natural languages can emerge from learning or use in isolation or whether (as predicted by Kirby et al., 2015) both learning and use are required to explain the structure of these linguistic systems.



Figure 1: Example Simple (left) and Complex (right) kin categorization systems. Generation is given vertically, relatedness is indicated by lines, males are on the left, females are on the right. While the Simple kinship system underspecifies gender (e.g. using the same label for "brother" and "sister"), the Complex system assigns the same label to very different meanings (e.g. paternal grandfather and maternal uncle).

- Berlin, B. & Kay, P. (1969). *Basic color terms: Their universality and evolution*. Berkeley: University of California Press.
- Carstensen, A., Xu, J., Smith, C. T., & Regier, T. (2015). Language evolution in the lab tends toward informative communication. In D. C. Noelle et al. (Eds.), *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.
- Greenberg, J. H. (1978). Generalizations about numeral systems. In J. H. Greenberg (Ed.), Universals of human language, volume 3: Word structure (pp. 249–295). Stanford: Stanford University Press
- Kemp, C., & Regier, T. (2012). Kinship Categories Across Languages Reflect General Communicative Principles. *Science*, 336, 1049–1054.
- Khetarpal, N., Majid, A., & Regier, T. (2009). Spatial terms reflect near-optimal spatial categories. In N. Taatgen et al. (Eds.), *Proceedings of the 31st Annual Meeting of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.
- Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, 141, 87–102.
- Regier, T., Kay, P., & Khetarpal, N. (2007). Color naming reflects optimal partitions of color space. *PNAS*, 104, 1436–41.
- Regier, T., Kemp, C., & Kay, P. (2015). Word meanings across languages support efficient communication. In B. MacWhinney & W. O'Grady (Eds.), *The Handbook of Language Emergence* (pp. 237–263). Hoboken, NJ: John Wiley & Sons.
- Murdock, G. P. (1970). Kin Term Patterns and Their Distribution. *Ethnology*, 9, 165–207.
- Xu, J., Dowman, M., & Griffiths, T. L. (2013). Cultural transmission results in convergence towards colour term universals. *Proceedings of the Royal Society B Biological Sciences*, 280, 20123073.

HOMO IMITANS IS HOMO LOQUENS. ON VOCAL IMITATION

JOANA ROSSELLÓ

joana.rossello@ub.edu Department of Catalan Philology and General Linguistics, Barcelona, Spain

Words are likely the most specifically linguistic ingredient of human language. Although conventional, rather than natural, they are, however, nature-culture hybrids since their acquisition depends on a biologically shared capacity for vocal imitation (VI). VI is present in vocal production learning, a scarce trait among animals that in primates is only shown in humans, the single homo loquens. By relating VI abilities, which are essential to build our huge vocabularies, to a biologically grounded vocal learning capacity in humans, the present proposal opposes to the view that VI is a developmental instance of a previous multimodal general imitative capacity (Donald, 1993) which would be central, innately given and unique to our species, thus seen as homo imitans (Meltzoff & Moore, 1997). Albeit recognizing the centrality of imitation in cooperative social interaction and cultural learning, I will argue to the contrary that VI has primacy and is co-opted to support general imitation. This view, besides solving the general neglect of VI in the literature on human imitation, may mediate between the two dominant psychological models of imitation, is parsimoniously explanatory, makes right empirical predictions regarding the signed modality and conditions such as autism and, finally, puts an inherently ostensive communication rather than action at the basis of human cognition. The Basic Vocal Imitation (BVI) approach might enlighten the debate between the Associative Sequence Learning (ASL) model (Catmur et al. 2009, Heyes 2016) and the Active Intermodal Matching (AIM) model (Meltzoff & Moore, 1997) underlying the homo imitans view. As a complement to ASL, BVI might provide a well-founded empirical basis for the ASL position that imitation mechanisms in humans are shared with animals. Besides, BVI circumvents the "correspondence problem" which goes with the mapping between "felt but unseen movements of the self with the seen but unfelt movements of the other"

(Heyes, 2016). Indeed, VI avoids the correspondence problem associated with the imitation of visual displays since one hears their own vocalizations as a result of the total feedback which is intrinsic and unique to vocal-auditory communication (Fitch, 2005). VI, on the other hand, is overtly developing by 8 months, as babbling unequivocally shows. In contrast, evidence for neo-natal non-vocal imitation is becoming increasingly weaker (Oostenbroek et al. 2016), which adds to the argument that the highly skilled imitation human abilities may be built on top of VI. Also overimitation, i.e., the imitation of actions that are causally irrelevant, is enlightened: it would be a uniquely human phenomenon among primates (Clay & Tennie, 2017) because humans are the only primates with vocal learning abilities and all vocal learners imitate *arbitrary* sequences of sounds, which in humans are in turn linked to arbitrary meanings in words.

If VI is basic, speech must be default. Sign is not default, it goes with deafness; and overlapping neural correlates for both speech and sign seen in native deaf signers arise from brain plasticity operating in sensory deprivation rather than being evidence for an alleged modality-independence of language. That the overlap lies in auditory areas in the brain is expected from this perspective. Regarding imitation in autism (ASD), Vivanti & Hamilton's (2014) review has shown that autistic individuals imitate less and worse than neurotypicals, mainly if the actions are sequential and not goal-directed (i.e. overimitation). Among explanatory proposals for this autistic deficit, Vivanti & Hamilton (2014) conclude that the sensorimotor deficit theory seems to be superior to the rest (e.g. 'failed direct self-other mapping' and other top-down proposals). Crucially, these authors do not mention either non-verbal autism (25% of ASD) or any deficit in VI, with the latter being inherent to the former. Yet, by putting an abnormal vocal learning capacity, with a consequent VI deficit, at the basis of ASD, the general imitative impairment as observed in the condition falls into place. The superiority of the sensorimotor deficit account also follows since sensorimotor integration is essential to VI. Finally, that nonverbal autistics can understand intentional actions and learn to communicate imperatively through images (PECS) rather than gestures, in which they are also impaired, strongly suggests that, contrary to the position in Scott-Phillips (2015), ostensive and declarative communication is primarily auditory and exapted to gestures.

References

Heyes, C. (2016). Homo imitans? Seven reasons why imitation couldn't possibly be associative. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1686)

- Meltzoff AN. (1988) The human infant as Homo imitans. In *Social learning: psychological and biological perspectives* (eds TR Zentall, BG Galef), 319–341. NJ: Erlbaum.
- Vivanti, G., & Hamilton, A. (2014). Imitation in ASD. Handbook of Autism and Pervasive Developmental Disorders, Vol. 1. Diagnosis, Development, and Brain Mechanisms, ed. By Volkmar et al., 278–302.

THE EFFECTS OF GOSSIP ON INTERPERSONAL TRUST DEPEND ON PROSOCIALITY OF THE GOSSIPERS

KONRAD RUDNICKI^{*1}, CHARLOTTE DE BACKER¹, and CAROLYN DECLERCK²

*Corresponding Author: konrad.rudnicki@uantwerpen.be
¹Department of Communication Science, University of Antwerp, Antwerp, Belgium
²Department of Management, University of Antwerp, Antwerp, Belgium

1. Evolutionary roots of gossip

Gossip is broadly defined as an evaluative information exchange about third parties (Foster, 2004). In 1998 Robin Dunbar developed a theory that gossip as a method of social grooming allowed humans to significantly expand their social group size (Dunbar, 1998). More specifically, gossip was proposed to boost ingroup trust. That idea was supported by studies showing that interpersonal gossip indeed promotes in-group cooperation (Boero et al., 2009; Feinberg et al., 2012; Piazza & Bering, 2008; Sommerfeld et al., 2008). Meanwhile, another theory was coined, that modern mass media are able to trigger mechanisms that evolved to deal with issues of interpersonal communication (Barkow, 1989). Because a modern subset of gossip – celebrity gossip, is a relatively new phenomenon it is not yet known if it is able to trigger the same effects as interpersonal gossip, and would its effects differ between face-to-face and online communication. This study is designed to test if celebrity gossip can enhance trust, just like interpersonal gossip does.

1.2. Prosociality

The social grooming theory of language states that gossip enables people to exchange information about the wrongdoings of others and punish the *free-riders* that exploit the group (Dunbar, 1998). Therefore, gossip can have a different effect on people depending on their attitude towards group living, since it protects those acting prosocially and threatens those who tend to be pro-self oriented. It was shown that prosocial individuals are intrinsically motivated to cooperate, and

the main determinants of trust for them are extraneous social cues signaling the trustworthiness of their interaction partner (Bogaert et al., 2008). Gossip is rich in social information and might create transparency by revealing reputation information about others. It has been shown that prosocially oriented people use reputation gossip concerning the wrongdoings of others to warn about potential *free-riders* (Feinberg et al., 2012). Therefore, we hypothesize that **celebrity gossip will increase trust of prosocial individuals (H**₁). Conversely, proselfs do not seek trust signals but situations with ambiguous social information that offer more opportunities to exploit others (Boone et al., 2010). Therefore, we hypothesize that **celebrity gossip will decrease trust of prosoli individuals (H**₂).

2. Methods

A 2x2x2 experimental design was used. Participants either gossiped about celebrities or performed a creativity task. Both tasks were performed either face-to-face or online with audio, but without video. Additionally, participants were classified as either prosocial or proself, based on Social Value Orientation questionnaire results (Van Lange et al., 1997). Trust was determined using an investment game with strategy method (detailed procedure in: Berg et al., 1995). 114 females participated in the experiment (M_{age} =21.96).

3. Results

Because participants interacted, the data was clustered by pairs of participants and analyzed as panel data in linear regression (Table 1).









4. Conclusion

In line with our hypotheses prosocials trusted their interaction partner more, whereas proselfs trusted their partner less after gossiping about celebrities. This provides support for the idea that gossip serves as an evoutionary mechanism to maintain the in-group cohesion which partially corroborates of the social grooming theory of language evolution (Dunbar, 1998). The fact that even modern celebrity gossip can have such an effect plays into the narrative that modern mass media are also able to trigger evolutionary mechanisms of interpersonal communication (Barkow, 1989). At the same time, the effect observed in our study was strictly dependent on social value orientation of the gossipers, which highlights how important it is to remember that evolutionary mechanisms can also be affected by individual differences (Nettle, 2007).

- Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and economic behavior, 10*(1), 122-142.
- Boero, R., Bravo, G., Castellani, M., & Squazzoni, F. (2009). Reputational cues in repeated trust games. *The Journal of Socio-Economics*, 38(6), 871-877.
- Bogaert, S., Boone, C., & Declerck, C. (2008). Social value orientation and cooperation in social dilemmas: A review and conceptual model. *British Journal of Social Psychology*, 47(3), 453-480.
- Boone, C., Declerck, C., & Kiyonari, T. (2010). Inducing cooperative behavior among proselfs versus prosocials: the moderating role of incentives and trust. *Journal of Conflict Resolution*, 54(5), 799-824.
- Dunbar, R., & Dunbar, R. I. M. (1998). Grooming, gossip, and the evolution of language. Harvard University Press.
- Feinberg, M., Willer, R., Stellar, J., & Keltner, D. (2012). The virtues of gossip: Reputational information sharing as prosocial behavior. *Journal of Personality and Social Psychology*, 102(5), 1015-1030.
- Nettle, D. (2007). Individual differences. In R. I. M. Dunbar & L. Barrett (Eds.), *Oxford handbook of evolutionary psychology* (pp. 479 – 489). Oxford, UK: Oxford University Press.
- Piazza, J., & Bering, J. M. (2008). Concerns about reputation via gossip promote generous allocations in an economic game. *Evolution and Human Behavior*, 29(3), 172-178.
- Sommerfeld, R. D., Krambeck, H. J., & Milinski, M. (2008). Multiple gossip statements and their effect on reputation and trustworthiness. *Proceedings of* the Royal Society of London B: Biological Sciences, 275(1650), 2529-2536.
- Van Lange, P. A., De Bruin, E., Otten, W., & Joireman, J. A. (1997). Development of prosocial, individualistic, and competitive orientations: theory and preliminary evidence. *Journal of Personality and Social psychology*, 73(4), 733-746.

DISTRIBUTION-BASED PREDICTION OF THE DEGREE OF GRAMMATICALIZATION FOR GERMAN PREPOSITIONS

DOMINIK SCHLECHTWEG* and SABINE SCHULTE IM WALDE

^{*}dominik.schlechtweg@gmx.de Institute for Natural Language Processing, University of Stuttgart

1. Introduction

Grammaticalization refers to the diachronic "development from lexical to grammatical forms, and from grammatical to even more grammatical forms" (Heine & Kuteva, 2007, p. 32). It is assumed to go along with *desemanticization*, a process of losing descriptive meaning (Heine, 2003; Bybee, 2015; Lehmann, 2015; Heine & Kuteva, 2007). For instance, the German noun Trotz 'defiance' acquired a less specific, grammatical meaning as preposition, cf. trotz des Sturms 'despite the storm', in which the original descriptive meaning is lost. The resulting prepositional meaning is semantically more general and often highly polysemous (Di Meola, 2014, cf. p. 134, 151). In addition, grammatical categories show a high degree of obligatoriness (Di Meola, 2014, cf. p. 39f.), i.e., they must be specified in a sentence (Lehmann, 2015, cf. p. 14). These three highlighted properties of grammaticalized expressions (generality, polysemy, obligatoriness) have a directly observable impact on their contextual distributions: they are used in a greater number of contexts (Weeds & Weir, 2003; Heine & Kuteva, 2007; Santus et al., 2014; Bybee, 2015; Schlechtweg et al., 2017). Together with the assumption that grammaticalization is a continuous process (Di Meola, 2014, cf. p. 68), these observations motivate our central hypothesis:

Hypothesis The degree of grammaticalization of an expression correlates with the unpredictability of its context words (contextual dispersion).

2. Method

In computational linguistics, a prominent corpus-based measure of contextual dispersion is word entropy (Hoffman et al., 2013; Santus et al., 2014). We exploit this measure in order to test our central hypothesis. First, we create a test set of German prepositions with different degrees of grammaticalization; we then (i) compute Spearman's rank-order correlation coefficient (ρ) between test set and word entropy scores, and (ii) use Average Precision (AP) to measure how well the scores distinguish between degrees.

For computing word entropy we induce a Distributional Semantic Model with window size 2 from a part of the SdeWaC corpus (Faaß & Eckart, 2013) with

approx. 230 million tokens. Low-frequency and functional words are deleted, and every token is replaced by its lemma plus POS-tag. For comparison, we also compute other quantitative measures of different aspects of contextual dispersion: word frequency and the number of context types.¹

3. Test Set

Di Meola (2014) distinguishes between (i) prepositions with the form of a content word (e.g., *trotz*), (ii) prepositions with the form of a syntactic structure (e.g., *am Rande*) and (iii) prepositions with the form of a function word (e.g., *vor*). Prepositions in (i) and (ii) show a low to medium degree of grammaticalization, while the ones in (iii) show a high degree (cf. p. 60). We focus on prepositions with the form of a PP from (ii), because Di Meola provides fine-grained distinctions, and (iii), to exploit a wide range of degrees. The final test set contains 206 prepositions with four degrees of grammaticalization (1: low – 4: high).²

4. Results

Table 1 shows that there is indeed a moderate correlation between entropy and the degree of grammaticalization, but frequency and the number of context types outperform entropy. Frequency has the highest overall correlation with grammaticalization: it is only .01 different to the correlation of context types (p = .6, two-tailed, Steiger's Z-test), but with .04 clearly different from entropy (p = .06). Frequency also distinguishes best between most of the degree levels; context types are generally comparable and in one case even the best predictor. Overall, the table clearly demonstrates that the more different the degrees of grammaticalization are, the better they are distinguished by the three measures.

	entropy	frequency	types
AP (degrees 1 vs. 2)	0.54	0.56	0.55
AP (degrees 1 vs. 3)	0.67	0.68	0.68
AP (degrees 1 vs. 4)	0.89	0.92	0.92
AP (degrees 2 vs. 3)	0.67	0.69	0.68
AP (degrees 2 vs. 4)	0.89	0.92	0.92
AP (degrees 3 vs. 4)	0.84	0.87	0.88
Spearman's ρ (rank)	0.42	0.46	0.45

 Table 1. Results for predicting degrees of grammaticalization.

Our findings contribute an empirical perspective to the relationship between grammaticalization and frequency, which has been discussed intensively (e.g., Di Meola, 2014, cf. p. 173) but not been investigated in a rigorous way, as done here.

¹Code: https://github.com/Garrafao/MetaphoricChange.

²The test set is provided together with the predicted measure scores as supplementary material.

- Bybee, J. L. (2015). *Language change*. Cambridge, United Kingdom: Cambridge University Press.
- Di Meola, C. (2014). *Die Grammatikalisierung deutscher Präpositionen* (2 ed.). Tübingen: Stauffenburg.
- Faaß, G., & Eckart, K. (2013). SdeWaC A corpus of parsable sentences from the web. In I. Gurevych, C. Biemann, & T. Zesch (Eds.), *Language processing* and knowledge in the web (Vol. 8105, p. 61-68). Springer Berlin Heidelberg.
- Heine, B. (2003). Grammaticalization. In *The Handbook of Historical Linguistics* (pp. 575–601). Oxford: Blackwell Publishing Ltd.
- Heine, B., & Kuteva, T. (2007). *The genesis of grammar: A reconstruction*. Oxford University Press.
- Hoffman, P., Lambon Ralph, M., & Rogers, T. (2013). Semantic diversity: A measure of semantic ambiguity based on variability in the contextual usage of words. *Behavior Research Methods*, 45(3), 718–730.
- Lehmann, C. (2015). Thoughts on grammaticalization. Language Science Press.
- Santus, E., Lenci, A., Lu, Q., & Schulte im Walde, S. (2014). Chasing hypernyms in vector spaces with entropy. In *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics* (pp. 38–42).
- Schlechtweg, D., Eckmann, S., Santus, E., Schulte im Walde, S., & Hole, D. (2017). German in flux: Detecting metaphoric change via word entropy. In Proceedings of the 21st Conference on Computational Natural Language Learning (pp. 354–367). Vancouver, Canada.
- Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, 87(2), 245–251.
- Weeds, J., & Weir, D. (2003). A general framework for distributional similarity. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, Sapporo, Japan (pp. 81–88).

COCKATIELS: A NOVEL ANIMAL MODEL FOR STUDYING THE EVOLUTION OF MUSIC AND LANGUAGE

YOSHIMASA SEKI

yoshimasa.seki@gmail.com Department of Psychology, Aichi University, Toyohashi, Japan

In view of the connections between music and language (Wallin et al., 2001), it is possible that research on the musical performance of animals could increase our understanding of the evolution of human language. Songbirds, such as zebra finches, which are widely used in lab experiments, are occasionally described as a model species for human musicality research. However, with some exceptions (e.g., bullfinches; Nicolai et al., 2014), members of this species seldom copy human music. Additionally, birdsong may differ substantially from human music (Araya-Salas, 2012). Meanwhile, spontaneous entrainment to various beats of human music by two parrots (Snowball and Alex) has been reported (Patel et al., 2009; Schachner et al., 2009). Thus, parrot species are potential and alternative candidates for the study of human music capabilities and evolutionary adaptation.

This study focuses on the cockatiel (*Nymphicus hollandicus*), a small cockatoo (parrot) species. While it is well known that some parrot species are excellent at imitating human speech, the cockatiel is characterized by its spontaneous imitation of human musical sounds using whistle-like vocal sounds, such as those heard in the *Mickey Mouse Club March*. This characteristic is well known by bird lovers (e.g. more than 2,500 search results were found for the key phrase ' "cockatiel" "mickey mouse march" ' on YouTube; as of Jan 2018); however, scientific studies evaluating vocal behavior in this species are limited.

I raised four cockatiel chicks that were exposed to the whistle sounds (an average of fewer than 5 times/day) of the *Mickey Mouse Club March*. At \sim 100 days post-hatch, three birds began to imitate the melody of the music. In the beginning, the birds would produce a short phrase, followed by another short phrase. The birds would then concatenate the short phrases and eventually complete the full piece with a slight alteration of the melody to a phrase. Two also

imitated several human words, but the other never did so, suggesting that the capability for vocal learning and that for imitating musical sounds may be independent of each other. The musical sounds were often produced immediately after the birds were left alone. Thus, although males seemed to be good at such mimicry, they may use these sounds to attract the attention of humans rather than as part of courtship behavior. Moreover, some birds showed a spontaneous "quick head shaking" during playback of human music. The timing of the behavior was analyzed to evaluate the rhythmic entrainment to the stimulus sounds. The results indicated a kind of regularity, but more evidence is needed to reach firm conclusions in this regard.

Acknowledgements

Supported by MEXT/JSPS KAKENHI JP25285198, JP17H06383 in #4903.

- Wallin, N. L., Merker, B., & Brown, S. (Eds.). (2001). The origins of music. Cambridge, MA: MIT press.
- Nicolai, J., Gundacker, C., Teeselink, K., & Güttinger, H. R. (2014). Human melody singing by bullfinches (Pyrrhula pyrrula) gives hints about a cognitive note sequence processing. *Animal Cognition*, 17, 143–155.
- Araya-Salas, M. (2012). Is birdsong music? Evaluating harmonic intervals in songs of a Neotropical songbird. *Animal Behaviour*, 84, 309–313.
- Patel, A. D., Iversen, J. R., Bregman, M. R., & Schulz, I. (2009). Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Current Biology*, 19, 827–830.
- Schachner, A., Brady, T. F., Pepperberg, I. M., & Hauser, M. D. (2009). Spontaneous motor entrainment to music in multiple vocal mimicking species. *Current Biology*, 19, 831–836.

WHAT MAKES A GRAMMAR DIFFICULT? EXPERIMENTAL EVIDENCE

ARTURS SEMENUKS^{*1}, and ALEKSANDRS BERDICEVSKIS²

*Corresponding Author: asemenuk@ucsd.edu

¹Department of Cognitive Science, University of California, San Diego, San Diego, USA ²Department of Linguistics and Philology, Uppsala University, Uppsala, Sweden

It is sometimes argued that linguistic structure can be affected by extralinguistic factors (Ladd, Roberts & Dediu, 2015). One specific hypothesis that has garnered a significant interest in the recent years links increase in the proportion of adult non-native (L2) learners in the population of a language to decrease in the morphological complexity of that language (Trudgill, 2011). Support for this view comes from qualitative typological analyses (Dahl, 2004; McWhorter, 2007; Trudgill, 2011), as well as correlational studies using both proxy (Lupyan & Dale, 2010) and direct measures of L2 proportion in the population (Bentz & Winter, 2013). This suggests that a greater L2 share in the population is associated with a higher probability of a language being analytic.

The exact details of how and why increased number of adult learners in the population can lead to morphological simplification, however, are not yet fully fleshed out. One hypothesis proposes that the simplification crucially depends on the cognitive differences between children and adults (Lupyan & Dale, 2010), as adults experience more difficulties learning morphology (DeKeyser, 2005). These difficulties, in turn, lead to the simplification of morphology.

McWhorter (2001) argues that grammars of creoles predominantly use analytic or isolating strategies, but an analytic morphosyntax by itself does not imply that a language had been affected by L2 speakers. However, his qualitative analyses do suggest that older non-creole languages have more grammatical distinctions compared to creoles (McWhorter, 2001). This begs the question of what exactly is more difficult for L2 speakers, and thus more likely to change in a contact situation: how many grammatical categories a language has or how complex the expression of these categories is (i.e. fusional is more difficult than agglutinative)?

To investigate this, we designed an artificial language (AL) learning experiment in which we manipulated the number of grammatical distinctions and the way they are expressed in the AL, and measured how these differences

affected the AL learning time. Native English speakers learned an AL that mapped onto intransitive English sentences with pronoun subjects, e.g. "I am walking". We separately manipulated (i) whether the pronoun was agglutinated to the verb and (ii) whether number was expressed fusionally, agglutinatively or not expressed at all in the AL, for a total of six conditions (Table 1). Thus four of the conditions expressed two grammatical categories and differed in the morphological complexity of their expression, and the two remaining conditions expressed only one category and differed in the complexity of its expression.

	Number &	Number & person	Number
	person fused	agglutinated	absent
Pronoun & verb	(1) potujota(2) lewujota	(1) pujota	(1) pujota
agglutinated		(2) pubijota	(2) pujota
Pronoun & verb	(1) potu jota(2) lewu jota	(1) pu jota	(1) pu jota
isolated		(2) pubi jota	(2) pu jota

 Table 1. Translations of example English sentences (1) "I am walking" and (2) "We are walking" in different experimental conditions.

The experiment continued until the participant made no mistakes in translating from English to AL. The number of rounds this took was measured. At the end, all of the participants translated six novel sentences, which served as a check for whether they truly learned the underlying generative rules of the language.

We recruited 48 participants per condition (288 total). ANOVAs showed that only the number of grammatical categories, but not the complexity of their expression, affected the learning time – ALs that did not encode number were learned faster. This held true both overall, as well as only in the data from participants that translated all novel sentences correctly (Figure 1). The results hint at the possibility that the difficulties in L2 learning are caused primarily by the number of grammatical categories, and not their morphological expression per se, which gives rise to the patterns observed in the studies discussed above.



Figure 1. Median numbers of learning blocks in different conditions for participants who were able to fully learn the language and generate novel sentences correctly.

- Bentz, C., & Winter, B. (2013). Languages with more second language learners tend to lose nominal case. *Language Dynamics and Change 3(1)*, 1–27.
- Dahl, Ö. (2004) *The growth and maintenance of linguistic complexity*. Amsterdam: John Benjamins Publishing.
- DeKeyser, R. M. (2005) What Makes Learning Second-Language Grammar Difficult? A Review of Issues. *Language Learning* 55, 1–25.
- Ladd, D. R., Seán G. R., & Dediu, D. (2015) Correlational studies in typological and historical linguistics. *Annual Review of Linguistics 1*, 221–241.
- Lupyan, G., & Dale, R. (2010) Language structure is partly determined by social structure. *PLoS ONE 5(1)*, e8559.
- McWhorter, J. H. (2001). The world's simplest grammars are creole grammars. *Linguistic typology*, 5(2), 125-66.
- McWhorter, J. H. (2007) Language interrupted: Signs of non-native acquisition in standard language grammars. Oxford: Oxford University Press.
- Trudgill, P. (2011) Sociolinguistic typology: Social determinants of linguistic complexity. Oxford: Oxford University Press.

EVIDENCE FOR HIGH FREQUENCY OF LINGUISTIC RARA IN EUROPEAN LANGUAGES

ARTURS SEMENUKS^{*1}

*Corresponding Author: asemenuk@ucsd.edu ¹ Department of Cognitive Science, University of California, San Diego, San Diego, USA

The concept of an absolute linguistic universal (ALU) – a feature that is shared or avoided by all languages – is important and useful in linguistic theory building. Indeed, presence of a truly absolute universal in a linguistic domain would impose an *absolute constraint* on the structures of explanatory theories for that domain, and thus highly likely have fundamental consequences for understanding cognitive and/or extracognitive pressures on linguistic structure, for the questions of how and why language evolved and evolves, and, ultimately, for specifying what are the defining features of language (Bickel, 2007; Piantadosi & Gibson, 2014).

However, although convincingly defending the existence of an ALU is a useful theoretical goal, some have put forward critical arguments for the position that currently there is no good evidence for the existence of many, if any, theoretically profound absolute universals (Dryer, 1998; Evans & Levinson, 2009). A single counterexample is enough to negate a proposed universal claim, which has been the fate of many of them, highlighting the variability in the potential structures that human language can obtain (Evans & Levinson, 2009). Furthermore, computational modeling suggests that due to sampling limitations rarely can typological investigations alone warrant considering a pattern an ALU, even for cases without exceptions among languages surveyed so far (Piantadosi & Gibson, 2014).

Because of this, it is crucial to better understand what leads to emergence of linguistic *rara* – features of languages that are extremely typologically uncommon and limited to only a handful of languages (Cysouw & Wohlgemuth, 2010). Such knowledge would provide assistance in finding counterevidence for ALUs by suggesting which languages a violation of a proposed ALU might be found in. The current work, in particular, aimed to investigate whether European

languages are overrepresented among the known languages with linguistic rara due to the researchers typically having more intimate knowledge of them.

To examine this, I web scraped the University of Konstantz electronic database of linguistic rara (Plank, 2006). The rara were matched with their corresponding area and genealogical stock using the AUTOTYP database (Nichols, Witzlack-Makarevich & Bickel, 2013). In order to safeguard against a stock being overrepresented due to retainment of a diachronically old and stable rare feature and simplify further analyses, only those rara that were reported to be found in a single language were left in the sample.

Statistical analyses of the data suggest that European languages are reported to contain a linguistic rarum that has been observed only in a single language more frequently then would be expected. Bootstrapping simulations also suggest that European languages contain unique rara more frequently than should be expected by chance (Figure 1).



Figure 1. The probabilities of European languages having different amounts of rara based on 10000 bootstrapping simulations. The dotted line denotes the observed amount of rara.

The results suggest two primary interpretations of the data. The first interpretation is that having a unique linguistic feature is more widespread among the world's languages than is usually assumed, and European languages stand out in the data due to implicit ethnocentrism of the researchers who are more familiar with the features of languages spoken in Europe, and so are more likely to spot a rarum in them. This could mean that the amount of ALUs is overestimated. The second interpretation suggests that European languages are indeed more unusual that the "cross-linguistic mean", which has implications for the sociological aspect of linguistic theory building. I will discuss the implications of both interpretations for the study of linguistic universals, and argue for the former interpretation.

- Bickel, B. (2007). Typology in the 21st century: major current developments. *Linguistic Typology*, 11(1), 239-251.
- Cysouw, M., & Wohlgemuth, J. (2010). The other end of universals: Theory and typology of rara. In: J. Wohlgemuth & M. Cysouw (Eds.), *Rethinking Universals. How Rarities Affect Linguistic Theories*. Berlin, New York: Mouton de Gruyter, 1–10.
- Dryer, M. S. (1998). Why statistical universals are better than absolute universals. *Chicago Linguistic Society 33: The Panels*, 123-145.
- Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and brain sciences*, 32(5), 429-448.
- Nichols, J., Witzlack-Makarevich, A., & Bickel, B. (2013). *The AUTOTYP genealogy and geography database: 2013 release.* Zurich: University of Zurich.
- Piantadosi, S. T., & Gibson, E. (2014). Quantitative standards for absolute linguistic universals. *Cognitive Science*, 38(4), 736-756.
- Plank, F. (ed.) (2006). *Das grammatische Raritätenkabinett*. Electronic database, online at: http://typo.uni-konstanz.de/rara/intro/

DIFFERENT PATTERNS OF ICONIC INFLUENCE IN THE CREATION VS. TRANSMISSION OF THE NICARAGUAN SIGN LANGUAGE LEXICON

ANN SENGHAS^{*1}, JENNIE PYERS², CAROLINE ZOLA¹ and CHARLOTTE QUINCOSES¹

*Corresponding Author: asenghas@barnard.edu ¹Department of Psychology, Barnard College, New York, NY, USA ²Department of Psychology, Wellesley College, New York, NY, USA

1. Background

Observations that iconicity diminishes over time in sign languages (Frishberg, 1975) pose a puzzle - why should something so evidently useful and functional decrease? Perhaps the pattern reflects differences between language-creation and language-transmission processes. If so, the effect might be mediated by the type of iconic representation: iconic signs can depict varied aspects of a referent object, such as how it is handled, or its shape, among other characteristics (Taub, 2001).

We ask, first, whether the prevalence and transparency of iconicity in sign languages results from its recruitment during language creation or during language learning; and second, whether language creation mechanisms favor the same types of iconicity that language-learning mechanisms do.

Existing languages are too mature to address these questions, but a newly emerging sign language in Nicaragua provides an opportunity to see the forces in play at language's earliest stages. Nicaraguan Sign Language (NSL) was created by deaf children and adolescents, starting with an initial cohort who arrived in a new special education school in the 1970s in Managua (Kegl et. al, 1999). Each subsequent age cohort to enter school developed the language further (Senghas & Coppola, 2001). By comparing the lexicons of different cohorts we can track the changes in iconicity over historical time, as NSL was transmitted to new learners.

2. Methods

As part of a larger study, we compared signs elicited in 2007 and 2017 from a first-cohort signer (35y in 2007) and a second-cohort signer (20y in 2007). From a set of 350 signs for everyday concepts, we selected 64, including 33 concepts rated as high-imageability (650-700) and 31 as low-imageability (200-450; Coltheart, 1981). Signs were categorized according to type of iconicity exhibited (see Fig. 1): either (a) pantomimic (e.g., embodying the referent itself, as in LION, or someone holding the referent, as in SPOON); (b) perceptual (e.g., representing the shape of a referent, as in TURTLE); or (c) arbitrary (e.g., non-iconic, as in TRUTH).



Figure 1. Examples of different types of iconicity in NSL lexical forms.

Three coders, blind to cohort and year, performed paired comparisons of each sign, determining whether any differences between the first- and second-cohort version indicated an increase, decrease, or no change in iconicity. Difference ratings considered both the transparency of iconicity (how easily a mapping could be inferred between the sign and its referent) and the number of articulators included in the iconic mapping (for example, if both the hand and mouth in LION mapped to the body of a lion, as opposed to only the hand). In this way, we compared how signs changed as they were coined and used by their creators over time, *vs.* passed down to new age cohorts of learners.

3. Results

The overwhelming majority of signs, particularly those for high-imageability concepts, exhibited some kind of iconicity. While the two cohorts' lexicons did not differ in the percentage of iconic signs, they did differ in their distribution of types, with the first-cohort signer producing slightly more pantomimic iconicity, and the second-cohort signer producing slightly more perceptual iconicity. Cross-cohort 2007 comparisons reveal that more signs decreased in iconicity (45%) than increased (14%), with pantomimic signs decreasing the most. Iconicity did not decrease overall from 2007 to 2017 (see Fig. 2).



Figure 2. Iconicity differences between cohorts in 2007 and 2017, by type of iconicity. A negative difference rating indicates that the second cohort sign was less iconic than the first cohort sign. In 2007, second cohort signs were less iconic than first, specifically for pantomimic signs. This suggests that pantomimic iconicity decreased as the second cohort acquired NSL in the 1980s. The pattern remains, though attenuated, in 2017.

4. Discussion

The high rate of pantomimic iconicity in the first-cohort lexicon reflects its high rate of recruitment during language creation. However, this type of iconicity also appears to be most subject to change, and was compromised during transmission to new learners. Accordingly, perceptual iconicity dominates in the second-cohort lexicon, suggesting that language-learning mechanisms may favor perceptual iconicity. We will consider differences between vertical and horizontal transmission processes, including visual, motoric, learning, and communication pressures, which may conspire to drive the asymmetry in iconicity type in sign language emergence.
Acknowledgements

Research funding provided by the James S. McDonnell Foundation (J.P.), Wellesley College Summer Science Center Research Program (J.P.), Barnard College Summer Research Institute (C.Z. & C.Q.), and National Institutes of Health/NIDCD grant R01 DC005407 (A.S.).

References

Coltheart, M. (1981). MRC Psycholinguistic Database User Manual: Version 1.

- Frishberg, N. (1975). Arbitrariness and iconicity: Historical change in American Sign Language. *Language*, 696-719.
- Kegl, J., Senghas, A., & Coppola, M. (1999). Creation through contact: Sign language emergence and sign language change in Nicaragua. In M. DeGraff (Ed.) Creolization, Diachrony and Development, 179–237.
- Senghas, A., & Coppola, M. (2001). Children creating language: How Nicaraguan Sign Language acquired a spatial grammar. *Psychological Science*, 12(4), 323-328.
- Taub, S. F. (2001). Language from the Body: Iconicity and Metaphor in American Sign Language. Cambridge: Cambridge University Press.

ABSTRACTION AND UNIVERSALS IN THEORIZING ABOUT LANGUAGE EVOLUTION

RICHARD SHILLCOCK

rcs@inf.ed.ac.uk

Philosophy, Psychology & Language Sciences, The University of Edinburgh, UK

Philosophical problems in the study of language evolution are introduced in connection with a range of different categories: totality, abstraction, universal, development, materiality. The approach which emerges contrasts with conventional theorizing in the field. The concept of 'concrete universal' is applied to reveal a continuity between the earliest necessary stages of language evolution and fully-developed contemporary language-use.

1. Philosophical problems in the language evolution field

Theorizing and experimentation come with often unexplored philosophical assumptions. I will address issues concerning the category of *universal* and associated categories, from the perspective of an explicit theory of knowledge. I will then develop a claim concerning continuity between the earliest stages in the evolution of language and fully-developed language use.

1.1. The totality

The scientific enterprise begins with the *totality* of the domain under investigation. For language evolution, it is the totality of our human ancestors collectively scraping an existence in the context of nature and culture, where the latter refers to habit, artefact, and ritual. (See MacWhinney, 2005, for a summary of some of the many factors that may have been important in language evolution.)

Critically, for the theory of knowledge employed here, the enterprise also *ends* with this same totality. As we will see, we will analyse this totality so as to identify component parts and characteristic mechanisms, but the exercise will end with us taking the 'return trip' away from partial views of the totality and finishing with the reconstituted totality. The scientific enterprise does not end with an analogy

between the totality and a representation couched in formal logic or in abstract mathematical terms. It ends back at the totality of the domain.

A second way in which cognitive scientists conventionally end up with partial views of the domain of language processing is that subdomains are considered in isolation: syllabicity, lexical stress, lexis, syntax, semantics, prosody, morphology, phonology and so on. The case of syntax is illustrative. The central phenomena-the 'fat' part of the distribution of sentence structures, the most frequent structures-are easily characterized, the challenge being to understand the peripheral phenomena, the 'long tail' of the distribution of sentence structures consisting often of only single instances in a large corpus (Steedman, pers. comm.). These rare items are caused by idiosyncratic interaction with the other subdomains. Hence the attention of formal linguists turns to such interaction-to morphosyntax, morphophonology and so on. The point I make here is that for abstract systems to interact there requires to be some form of *mediation*. The more abstract systems that interact, the likelier it is that the mediator will be something material, something real we can observe and manipulate. A material mediator is richer in potential 'points of contact' and is maximally applicable to new forms of interaction. (We return below to this issue.)

A third way in which cognitive science approaches to language processing remain at an abstract level is the conventional goal—implicit in the typical approach to universals—of generalizing across all languages, of characterizing Language with a capital L. The overall tendency is for 'explanation' to apply to Language and 'description' to apply to specific languages. The approach pursued here is concerned with material explanation. In the second half of the paper we address the case of English. This move is not due to the convenience that more psycholinguistic research exists for English than for any other language. Rather it is a recognition that the approach should be applied to material mechanisms in real and typical cases. Any real language is a suitable domain of study. The best explanation for the development of English cannot be disproved or improved by data from an unconnected language. We return below to the question of other languages.

1.2. Abstraction

To make the initial totality tractable, we need to *abstract* from it. The type of abstraction required is an 'Aristotelian abstraction', in which we take away (in the mind of the analyst) aspects of the totality (*cf.* Jones, 2005). This is fundamentally different from 'Galilean abstraction' and the example of the frictionless plane that is typically cited (cf. Weisberg, 2007). This latter approach will mislead us. First,

because rather than taking something away from the domain, it is truer to say that Galilean abstraction *replaces* something; a plane that involves real friction has been replaced by an abstract plane that stands for that which is similar across all planes of varying friction. Second, experience shows that researchers that employ Galilean abstraction frequently do not cash in the promise to return to the complete domain, preferring instead to stay with the abstract view of the domain.

1.3. A universal

What are we expecting to find by successively taking away material aspects of the domain? Our goal is to identify the simplest entity that itself *simultaneously exhibits and explains* the behaviours we see in the domain.

Vygotsky (1934/86) termed such an entity the 'unit of analysis'. His example was the investigation of the fluid properties of water. The relevant unit is the water molecule. Its constituent hydrogen and oxygen atoms are too small to exhibit the relevant behaviours—van der Waals force, *etc.*—that lead us to an understanding of the fluidity of water.

Such an entity is a universal. It unifies many things—the definition of a universal. It participates *as itself* in the relevant domain. We required to *discover* it in the domain. It *mediates* everything else in the domain, in which it is pervasive. It is a *material* entity—it is not going to be 'disproved' or wished away by new data.

Such a universal is different in all these respects from the type of universal that cognitive scientists are used to employing. Take 'verb'. This is an entity that needed to be created, not discovered. The category 'verb' does not participate as itself in the domain (although we may initially think so, because we tend to turn formal linguistic categories into processing categories). Verb is certainly a useful category, though. It captures generalizations about the domain, in the form of *ordered relations* between itself and other such abstract universals (syntax being the paradigm example of such relationships). But this is the limit of the category 'verb', combined with the fact that it may be eclipsed by new data (such as the weakness of the formal distinction between nouns and verbs in the Salish languages; Jelinek, 1995).

This less conventional universal has a long philosophical history (Shillcock, 2014) in which it is termed the *concrete universal*, signaling its role in connecting other entities.

1.4. Explanation

In this materialist theory of knowledge, explanation consists in having identified the concrete universal within the domain by a process of successively abstracting away the peripheral aspects of the domain and in then being able to reverse the process by successively taking this or that aspect of context and demonstrating the *necessary* implications it has for the role of the concrete universal within the domain, until the totality of the domain has been recreated.

If at any point in the movement between the concrete universal and the totality, we can say that adding this aspect of context causes this to happen, and subtracting this other aspect of context causes this other thing to happen, then we have achieved an explanation of how the domain works. We will have understood the 'logic' of the domain—not in the sense of some abstract formalism, but in terms of the material role of the concrete universal.

Another example. The *stem cell* in the domain of human anatomy graphically represents Vygotsky's notion of the unit of analysis or *cell-form* of a domain. The stem cell participates as itself in the domain and exhibits all of the basic characteristics of the other cells of the body. The stem cell is pervasive, it mediates the life of every other cell in that all other cells are derived from stem cells. If we can specify every aspect of the contexts that *necessarily* cause a stem cell to turn into a neuron, a skin cell, a lung cell, and so on, and we can characterize the effects on the stem cell at each point in the emergence of the organs of the body, then we will have explained human anatomy.

1.5. The genetic approach to studying language

The approach outlined above is essentially a *genetic* approach (*cf.* Vygotsky, 1934/1986). A domain is understood by exploring how it came into being. In this sense, the study of the development of language (phylogenetic and ontogenetic) is the best way to understand language.

At the same time, the current necessity of beginning with the developed form of language (as opposed to having access to some still-existing earlier stage) is actually a *privileged* viewpoint, in that we can identify potentially insignificant aspects of the possible beginnings of language that go on to become crucial in the developed form. Similarly, we can put into perspective potentially important aspects that have little role in the fully developed form.

1.6. Materiality of a concrete universal

It is the material nature of the concrete universal that goes most against the grain of conventional cognitive science. The credo of 'Marr's levels' (Marr, 1982) has as its highest level the computational specification of the problem—the laws of arithmetic, in the typical example. Marr abstracted from the totality of the processing and did not then perform the 'return journey' to explain how any essence of the processing might emerge in the first place and develop into the observed role in the full context.

In the present approach to explanation, we require both *individuality* and *generality*. We require to know exactly what is going on throughout the domain of study. Nothing can be absolutely individual or absolutely general. It is in the dialectic between these two goals that the concrete universal emerges—an entity as detailed as possible and with as pervasive a mediating role as possible. These requirements necessitate a *material* entity as the concrete universal in the domain.

2. Applying the analysis

Below, the approach sketched in 1 is applied to the evolutionary emergence of language. There is necessarily a continuum of development with the processing of current English. How can our understanding of the latter inform our exploration of language evolution?

The first move we must make is to define the domains with which we are concerned. Depending on the content of the domain, a different concrete universal will apply; the 'essence' of what goes on in the domain will be different. We can consider three domains: 'general human cognition', 'communication', 'contemporary English'.

2.1. The domain of general human cognition

The theoretical positions of this paper suggest that *hemisphericity* is central to the emergence of general human cognition (Shillcock, Thomas, & Bailes, *submitted*). Hemisphericity is the relative encapsulation, differentiation and autonomy of the cerebral hemispheres; the degree of hemisphericity distinguishes the human brain from the even our closest primate relatives. The material requirement of the hemispheres to coordinate themselves is the relevant concrete universal. This coordination necessitates each hemisphere modelling the other in the context of body and world, effectively 'parsing the world' and thereby generating tools for cognition. 'Predictive Processing' (Friston, 2005; Clark, 2015) between the hemispheres is seen as playing an important role in this mutual modelling, but

each hemisphere effectively uses the other hemisphere as a tool. We return to the role of the hemispheres below.

2.2. The domain of communication

What is the concrete universal in the domain of communication? Perhaps the best candidate for the 'cell-form' in this domain is the addresser materially re-orienting the addressee between old and new information: "you already know this" *vs.* "this is news". This distinction is the very basis of communication. The use of 'orienting' here recalls the orienting response (Sokolov, Spinks, Näätänen, & Lyytinen, 2002) to an external stimulus and a similar response to an 'internal stimulus'.

Language and cognition are heavily-tilled fields and we can expect to encounter familiar distinctions; in this case, Sperber and Wilson's (1995) 'relevance'. What is new here is situating the distinctions within the deeper particular philosophical and psycholinguistic contexts.

We should expect to be able to identify some material orienting activity by the addresser. Fundamentally, the *addresser uses the addressee as a tool* to achieve some purpose, as in physically steering them to participate in joint activity. Physical contact, non-contact gesture, and speech sounds are all equivalent in the sense that they reorient the addressee between old and new information. They naturally differ in their directness and sophistication, spoken language being the most developed form of re-orienting the addressee.

If we limit our considerations to speech, then the minimal vocal gesture is that produced by the vocal apparatus approximating a simple tube—a spectrally diffuse schwa-like sound (*cf.* Harris, 1995). We can identify such a minimal reorienting activity with the contemporary cross-linguistic vocal gesture identified by Dingemanse, Torreira and Enfield (2013) as 'other-initiated repair'. In this joint activity, the interlocutors negotiate the relationship between old and new information.

2.3. The domain of contemporary English

We have identified a material universal—the activity of producing a schwa-like sound—as playing the role of *re-orienting the addressee between old and new information*. The critical move we now make is to continue exploring the role of this concrete universal beyond the simple routine of repair initiation and repair solution.

Vocal activity approximating a schwa sound accounts for some 10% of conversational English (Fry, 1947). It fits the bill for a concrete universal for contemporary English by virtue of being pervasive within the domain.

Schwa-like activity in spoken English has a still stronger claim to be the relevant concrete universal. It mediates every other entity in the domain. The speech activity that is conventionally labelled schwa has an idiosyncratic relationship with each of the formal subdomains explored in contemporary spoken English, from articulatory phonetics, through phonology, lexical stress, syllabic structure, the functor/contentive distinction, syntax (by way of projection from functional heads), and even word meaning.

Consider just one example of schwa-activity from the above partial list. Schwa-activity is a reliable cross-linguistic indicator of functorhood (Shi, Morgan, & Allopenna, 1998). A critical development in syntactic theory occurred with the identification of the functor as the syntactic head of the phrase as opposed to the noun. This change was very productive in revealing systematicity in syntax. It is one instance in which schwa-activity constitutes the material bridgehead into a formal subsystem of language (syntax).

Schwa-activity constitutes a similar material bridgehead into the other formal subsystems of contemporary English. It is contained in each subsystem, constrained in very specific ways, and is a material way in which those subsystems articulate one with another.

The claim here is emphatically *not* that developed linguistic behaviour *reduces* to schwa-activity. It is that schwa-activity in its primitive form of negotiating the relationship between old and new information is the basic cell-form that can characterize and explain more fully developed linguistic behaviour.

Thus, the simple form of schwa-activity may be elaborated phonologically (adding or subtracting behaviours corresponding to features) to provide the wherewithal to make more detailed linguistic references. It mediates subsystems such as syllabicity and lexical stress, which in turn provide more structure and expressive potential.

Critically, schwa-activity in contemporary English signals old information, a return to the shared world, a retreat from new information (see, *e.g.*, Fowler & Housum, 1987). It occurs in functors signalling known aspects of the discourse such as tense and number. It occurs in parts of words beyond the uniqueness point.

2.4. The constitutive role of schwa-activity in processing

The claim was rehearsed above that the necessity of hemispheric coordination was the concrete universal that generated distinctively human cognition. This occurs by mutual hemispheric modelling occurring partly by direct callosal interaction, partly by predictive processing, and partly by the tool-mediated interaction of the contralaterally-controlled manipulators.

The frequent occurrence of schwa-like activity in speech marshals the hemispherically-distributed processing preferences for prosody, phonology, different aspects of syntax, differently parameterized semantics, face processing, lexis and so on. It reinforces their mutual relationships.

This approach to a theory of knowledge thus takes us from the emergence of communication within the species, where the concrete universals suggested—the necessity of hemispheric communication and the orientation between old and new information—will be fundamental and it delivers us to the more idiosyncratic constraints of the cultural evolution of a specific language, where the suggested concrete universal—schwa-like activity—works for English, but where some analogous speech activity may be required for different languages.

3. Conclusions

We have introduced a universal, not conventionally recognized as such, but assumed by a materialist theory of knowledge—the *concrete universal*. We have applied this theoretical tool to the domains relevant to the early stages of language evolution—'general human cognition' and 'communication'—and we have identified the material 'cell-form' of these domains—schwa-like activity. We have then crucially suggested a continuity between these early aspects of language evolution and the continuing development of linguistic activity and structure. Our domain of application has been contemporary English, but the prediction is that an analogous analysis can be made for any other real and typical language.

The approach we have followed firmly orients researchers towards the material richness of language behaviours and encourages them to interrogate their models—from box-and-arrow diagrams to computational implementations—as to the precise status of each entity in the model. Experience shows that the serious limitations of abstract universals are overlooked and that the merits of concrete universals are not appreciated.

The goal has been to explore the 'logic' by which fully-fledged linguistic activity works and to draw a line of material continuity between such fully-developed processing and the earliest stages of language evolution.

References

Clark, A. (2015). Surfing uncertainty: Prediction, action, and the embodied mind. Oxford University Press.

- Dingemanse, M., Torreira, F., & Enfield, N. J. (2013). Is "Huh?" a Universal Word? Conversational Infrastructure and the Convergent Evolution of Linguistic Items. *PloS one*, 8(11), e78273.
- Fowler, C. A., & Housum, J. (1987). Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language*, 26(5), 489-504.
- Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 360(1456), 815-836.
- Fry, D.B. (1947). The frequency of occurrence of speech sounds in Southern English. Archives Néerlandaises de Phonétique Expérimentale, 20, 103-106.
- Harris, J. (2005). Vowel reduction as information loss. In P. Carr, J. Durand, & C. J. Ewen (eds.), *Headhood, Elements, Specification and Contrastivity*, pp. 119-132. Amsterdam: Benjamins.
- Jelinek, E. (1995) Quantification in Straits Salish. In: Quantification in natural languages, ed. E. Bach, E. Jelinek, A. Kratzer & B. Partee, pp. 487–540. Kluwer.
- Jones, M. (2005). Idealization and abstraction: A framework. In M. Jones and N. Cartwright (eds.), *Idealization XII: Correcting the Model: Idealization and Abstraction in the Sciences* (New York: Rodopi), 173-217.
- MacWhinney, B. (2005). Language evolution and human development. In Bjorklund, D. and Pellegrini, A. (Eds.). Origins of the Social Mind: Evolutionary Psychology and Child Development, New York, Guildford Press.
- Marr, D., & Vision, A. (1982). A computational investigation into the human representation and processing of visual information. *WH San Francisco: Freeman and Company*, *I*(2).
- Shi, R., Morgan, J. L., & Allopenna, P. (1998). Phonological and acoustic bases for earliest grammatical category assignment: A cross-linguistic perspective. *Journal of child language*, 25(1), 169-201.
- Shillcock, R. (2014). The Concrete Universal and Cognitive Science. *Axiomathes*, 24(1), 63-80.
- Shillcock, R., Thomas, J., & Rachael Bailes, R. (*submitted*). Mirror Neurons, Prediction and Hemispheric Coordination; The Prioritizing of Intersubjectivity over 'Intrasubjectivity'.
- Sokolov, E. N., Spinks, J. A., Näätänen, R., & Lyytinen, H. (2002). The orienting response in information processing. Lawrence Erlbaum Associates Publishers.
- Sperber, D., & Wilson, D. (1995). Relevance: Communication and cognition (2nd. ed.). Oxford, England: Blackwell.
- Vygotsky, L. (1934/1986). Thought and language. Cambridge, MA: MIT Press.
- Weisberg, M. (2007). Three kinds of idealization. The Journal of Philosophy, 104(12), 639-659.

STRUCTURAL ALIGNMENT IN CROSS-MODAL PRIMING: LINGUISTIC REPRESENTATION IS SHARED BETWEEN GESTURE AND SPEECH

JOLEANA SHURLEY¹, MARIEKE SCHOUWSTRA², and MARTIN J. PICKERING^{3 3}

^{1, 2} Centre for Language Evolution, University of Edinburgh, UK ¹ s1470688@ed.ac.uk, ² Marieke.Schouwstra@ed.ac.uk

³ Department of Psychology, University of Edinburgh, UK martin.pickering@ed.ac.uk

There are robust findings that support a theory of interactive alignment, which indicate that communicative partners engage in various levels of synchrony during dialogue, both in the modality of speech as well as in gesture (Branigan et al., 2000; Schouwstra & de Swart, 2014; Christiansen et al., 2016). Interactive alignment is largely a consequence of priming (Pickering & Garrod, 2004). Recently it has been found that priming plays a role in artificial language change towards maximally regular and highly predictable linguistic systems (Feher et al., 2016), and in the emergence of linguistic conventions in silent gesture studies (Schouwstra et al., 2016). Silent gesture experiments, which use hearing participants, indicate that an individual's choice of constituent order (e.g., SOV) is primed by their communicative partner's own syntax. To further investigate the extent of interactive alignment within cross-modal communication, from gesture string comprehension to verbal speech production.

We used a picture-matching game to test structural alignment between the syntax used in gesture strings and the word order produced in speech. A confederate to the study gestured events depicted in a set of pictures (see Figure 1), and the participant chose the correct card out of an array presented on-screen. Subsequently, the participant *verbally* described their own set of target cards to the gesturer. Crucially, the gesturer followed a script that systematically varied in the syntactic structure of the dative alternation (depicting PO, prepositional-object, or DO, double-object sentences) when manually describing the set of cards. Participants were led to believe they were interacting with the gesturer by sending recordings via a network connection. In fact, the participants were viewing pre-recorded videos of gestures.

Figure 1: Experimental item drawn from the gesturer's description set of virtual cards, corresponding to the double-object dative sentence: "[A/The] boxer *shows* [a/the] ballerina [a/the] book."



The verbal descriptions gathered from the participant served as the target responses used in analysis. Table 1 reports the proportion of target responses in the different conditions. Half the participants described cards that were the same as those actions depicted in the prime ('throw', 'sell', 'show', 'give'), and the other half described cards that were mismatched, or different. We investigated whether using the same lexical item and a similar syntax provided a 'lexical boost', whereby participants produced more linguistically aligned target responses than if they were primed by similar syntax alone (Branigan et al, 2000). To obscure the nature of the study, as well as priming methods, these items were interspersed with filler items, which depicted simple transitive sentences (where only one direct object is obligatory).

Table 1: Percentage of target responses in each of the four conditions				
Verb type	Prime Description	Target description		
		РО	DO	
Same	PO	0.92	0.08	
	DO	0.79	0.21	
Different	PO	0.95	0.05	

.

PO: prepositional-object; DO: double-object

We ran a linear mixed effects model looking at the proportion of DO produced, entering prime order and sameness of verb as fixed effects and participant and verb target as random effects. Overall, participants had a preference for PO over DO, which is reflected in the intercept (beta=-2.07, SE=.39, p<.001). In trials following a DO prime, they were significantly more likely to produce DO (beta=0.915, SE=.41, p<.05). No interaction was observed between prime order and sameness of verb (beta=0.151, SE=.55, p=.78).

Our results illustrate that structural properties of silent gesture influence the syntactic structure of verbal descriptions, i.e. showing evidence of cross-modal priming. This is interesting because silent gesture, when used in production, has been shown to by-pass native language rules (Goldin-Meadow et al., 2008), and can be seen as the result of an improvisation process, reflecting cognitive biases of the gesturer (Schouwstra & de Swart, 2014). The fact that spoken English can be primed by this mode of communication shows the far-reaching effect of silent gesture, and its potential role in emerging language situations (Kelly et al., 2008). This interpretation is strengthened by the fact that no lexical boost effect was observed in this experiment: participants did not simply 'translate' the gesture strings into their native language.

- Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition*, 75(2), B13-B25.
- Fehér, O., Wonnacott, E., & Smith, K. (2016). Structural priming in artificial languages and the regularisation of unpredictable variation. *Journal of Memory and Language*, 91, 158-180.
- Goldin-Meadow, S., So, W. C., Özyürek, A., & Mylander, C. (2008). The natural order of events: How speakers of different languages represent events nonverbally. *Proceed*ings of the National Academy of Sciences, 105(27), 9163-9168.
- Kelly, S., Manning, S., Rodak, S. (2008). Gesture gives a hand to language and learning: Perspectives from cognitive neuroscience, developmental psychology and education. *Language and Linguistics Compass*, 2(4), 569-588.
- Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral* and brain sciences, 27(2), 169-190.
- Schouwstra, M., Smith, K., & Kirby, S. (2016, March). From natural order to convention in silent gesture. In The Evolution of Language: Proceedings of the 11th International Conference (EVOLANG XI). EVOLANG XI, New Orleans, United States 20, 24.
- Schouwstra, M., & de Swart, H. (2014). The semantic origins of word order. *Cognition*, 131(3), 431-436.
- Christensen, P., Fusaroli, R., & Tylén, K. (2016). Environmental constraints shaping constituent order in emerging communication systems: Structural iconicity, interactive alignment and conventionalization. *Cognition*, 146, 67-80.

THE INFORMATIVE VALUE OF COMMUNICATOR DISAGREEMENT IN GREAT APES FOR THE STUDY OF THE EVOLUTION OF HUMAN INTENTIONAL COMMUNICATION

CHRISTINE SIEVERS^{*1,2}, THIBAUD GRUBER^{3,4}, and KLAUS ZUBERBÜHLER³

*Corresponding Author: christine.sievers@unibas.ch ¹Department of Philosophy and Media Studies, University of Basel, Basel, Switzerland ²Department of Comparative Cognition, University of Neuchâtel, Neuchâtel, Switzerland ³Department of Zoology, University of Oxford, Oxford, UK ⁴Swiss Center for Affective Science, University of Geneva, Switzerland

Human intentional communication is a complex back- and- forth between communicators, with speakers providing overtly information to hearers and hearers aiming to understand the transmitted information and responding accordingly, in line with the provided information and their own intentions. In regards to how such a complex form of communicative interaction has evolved from simpler forms of communication, nonhuman ape communication can serve as а model. Intentional great communication in nonhuman primates though is often analyzed solely from the perspective of the signaler. Signalers are supposed to stop producing the signal when their goal is fulfilled and they should display persistence or elaboration behavior in case their goal is not fulfilled (Townsend et al., 2016; Leavens et al., 2005). A recipient's reaction to a possible intentional production from the signaler is then largely ignored. In this framework, situations where no information is openly communicated by the signaler or where recipients make decisions independent of the signaler's potential goal and the signal's intended meaning may qualify as instances of intentional communication. Following from that, very simple forms of interaction may be labeled intentional communication, even though these interactions do not have much in common with human intentional communication. The key to providing a more informative comparison is to explicitly focus on recipient's reactions and signaler's responses in the species supposed to serve as a model. Particularly, cases where producer and recipient disagree are informative because both actors

can provide clues of their disagreement by indicating a diverging goal or displaying dissatisfaction. Such disagreement may amount to a back-andforth between communicators communicating their respective goals, just as it is the case for human communication. Human communicators tend to engage in a back-and-forth in elaborating openly and extensively on their intentions especially in cases in which the communicators disagree on something. If we found such instances in nonhuman primates, it may first be concluded that nonhuman primates can communicate intentionally in a true sense of intentional communication, that is overtly; second, one could subsequently describe the differences between human and nonhuman intentional communication. To address this, we looked at wild chimpanzees as a model species to find instances of communicative disagreement. We video-recorded travel initiations in wild chimpanzees (Pan troglodytes schweinfurthii) in Budongo Forest, Uganda, using focal animal sampling over a six-month study. We were particularly interested in instances where one individual had the goal of leaving with a recipient, but the recipient did not follow and therefore both engaged in a back-andforth interaction. We found 21 occurrences of such conflicts of interest out of 283 travel initiations (7%) in various contexts including family travel and courtship. To analyze these instances, we studied vocalizations, gestures, the orientation of the body, the moves and gazes of the partners towards the travel direction and each other, and whether they displayed aggression or extensive waiting. Our data show that wild chimpanzees are capable of communicating and negotiating diverging goals during travel to reach a common outcome, though this remains rare. In conclusion, we documented clear instances of disagreement between chimpanzees and their behavior in attempt to win over the other individual, demonstrating how especially vocalizations play an important role, with up to 32 communicative back-and-forths between chimpanzees, suggesting similarities to human conversation, not just with regard to its turn-taking structure but also in regards to its overt nature. Looking at the distribution of behaviors between initiators and potential followers in all contexts, initiators produced more than double the no. of signals and behaviors. Furthermore, negotiations had the longest duration in consortship contexts, with potential followers displaying significantly more communicative behavior than in other group contexts.

Acknowledgements

This research was funded by the Swiss National Science Foundation.

- Leavens, D. A., Russell, J. L. & Hopkins W. D. (2005). Intentionality as measured in the persistence and elaboration of communication by chimpanzees (Pan troglodytes), *Child Development*, 76, 291-306.
- Townsend, S. W., Koski, S., Byrne, R., Slocombe, K., Bickel, B., Braga Goncalves, I., Boeckle, M., Burkart, J., Flower, T., Gaunet, F., Glock, H., Gruber, T., Jansen, D., Liebal, K., Linke, A., Miklosi, A., Moore, R., van Schaik, C., Stoll, S., Vail, A., Waller, B., Wild, M., Zuberbühler, K. & Manser, M. (2016). Exorcising Grice's ghost: an empirical approach to studying intentional communication in animals, *Biological Reviews, doi: 10.1111/brv.12289. [Epub ahead of print].*

'GIVE IT TO ME'/ 'GIVE IT TO HER': THE DEVELOPMENT OF A RELATIONAL TERM IN HOMESIGN AND ENGLISH

Catriona Silvey^{*1} and Susan Goldin-Meadow²

^{*}Corresponding Author: c.silvey@ucl.ac.uk ¹Division of Psychology and Language Sciences, University College London, London, UK ²Departments of Psychology and Comparative Human Development, University of Chicago, Chicago, USA

In typical language acquisition, evidence from naturalistic and experimental studies suggests that the meanings of relational terms such as verbs are harder for children to learn than the meanings of nouns (Gentner, 1982; Gleitman, Cassidy, Nappa, Papafragou, & Trueswell, 2005). Children's eventual acquisition of verb meanings is explained as being supported by other aspects of the linguistic system. By repeatedly highlighting relations across contexts (Loewenstein & Gentner, 2005) or bootstrapping meaning through syntax (Gleitman et al., 2005), the structure of language as a whole is argued to be crucial to children's acquisition of relational terms. However, from a language evolution perspective, this explanation is not entirely satisfactory: if a fully developed linguistic system is required to learn these meanings, then what is their evolutionary origin?

One source of evidence to address this question is homesign. Homesigners are deaf children raised by hearing parents who do not expose them to a sign language. These children are unable to access spoken language input, and their parents' gestural input is unsystematic; however, the children spontaneously invent gestural systems for communication, referred to as homesign. These systems have many of the features of natural languages (Goldin-Meadow, 2003), including a distinction between nouns and verbs (Goldin-Meadow, Butcher, Mylander, & Dodge, 1994). Homesign is a crucial and hitherto overlooked source of evidence in the debate about the acquisition of relational terms. By examining the similarities and differences between the parent's gestural input and the child's homesign, we can determine which relational concepts are developed and generalized on the basis of evidence from input, versus factors internal to the child.

We focus here on the relational concept lexicalized by the English word 'give', and specifically on the role of the recipient. We compare the expression of this concept in the productions of one homesigner, David, observed over 4 sessions from the ages of 34 to 39 months, and of 15 typically developing children acquiring English, observed over 12 sessions from the ages of 14 to 58 months.

From his first session, David produces a gesture (GIVE), which has a stable form (Goldin-Meadow et al., 1994) and, like the English verb 'give', appears to be a three-place predicate involving an agent, a patient and a recipient (Goldin-Meadow et al., 1994). Initially, during the 3 sessions from 34 to 36 months, the recipients in David's uses of GIVE are restricted to himself (47 tokens) or locations adjacent to him (2 tokens). Only later, in the 39 month session, does he use GIVE with another person as a recipient (2 of 9 tokens). There are two potential explanations for this change. One is that it is due to maturational or motivational factors. Another possibility is that this an effect of input: during the previous session, when David was 36 months old, his mother gestured GIVE and specified David's sister as a recipient. This sequence is compatible with the hypothesis that David's GIVE gesture initially has a meaning closer to 'give-to-me/put-near-me', which he then generalizes on the basis of maternal input that suggests a broader range of uses. However, before drawing this conclusion, it is crucial to look at children exposed to typical natural language input for comparison.

We analyzed early productions of 'give' by 15 typically developing children acquiring English. During the first observational session in which they produced 'give' (average 3 tokens), all but 2 of the 15 children used it in a similarly restricted way to David, i.e., with only the self as recipient. All but 1 child later produced 'give' with other recipients in subsequent sessions. To determine whether variation in children's onset of using 'give' with other recipients could be explained by input, we analyzed parents' use of 'give' (average 22 tokens) prior to each child's first production of the verb. All parents produced 'give' with other recipients during sessions prior to their child's first production of the verb. Thus, children's uses of 'give' are initially restricted despite evidence for generalization. Furthermore, the amount or proportion of parental uses of 'give' with other recipients does not predict children's age at first using 'give' with other recipients, suggesting that input is not the decisive factor for children's productive generalization; motivational or other factors may be involved. Indeed, David's age of onset of GIVE with other recipients (39 months) falls close to the median age of onset in English-acquiring children (38 months). Previous work showed that children's verbs encode their own actions before they encode the actions of others (Huttenlocher, Smiley, & Charney, 1983). Our results suggest that rather than always focusing on the self as actor, children may initially use the role most salient to them (recipient, in the case of GIVE) as a 'way in' to the argument structure of a verb.

The factors that influence the development of GIVE/'give' in production may be relatively independent of language input. However, other relational terms, such as mental state verbs, are notably absent from homesign. This suggests an evolutionary story where some relational terms can develop without language input, whereas others need the support of a full linguistic system. Future work should investigate the factors that determine which relational terms fall into which category, and the processes by which more language-dependent relational terms emerge.

- Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In *Language development, volume 2: Language, thought and culture* (pp. 301–334). Hillsdale, NJ: Lawrence Erlbaum.
- Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard words. *Language Learning and Development*, 1(1), 23–64.
- Goldin-Meadow, S. (2003). *The resilience of language: What gesture creation in deaf children can tell us about how all children learn language.* New York: Psychology Press.
- Goldin-Meadow, S., Butcher, C., Mylander, C., & Dodge, M. (1994). Nouns and verbs in a self-styled gesture system: what's in a name? *Cognitive Psychology*, 27, 259–319.
- Huttenlocher, J., Smiley, P., & Charney, R. (1983). Emergence of action categories in the child: Evidence from verb meanings. *Psychological Review*, *90*(1), 72–93.
- Loewenstein, J., & Gentner, D. (2005). Relational language and the development of relational mapping. *Cognitive Psychology*, *50*(4), 315–53.

REFERENTIAL GESTURES ARE NOT UBIQUITOUS IN WILD CHIMPANZEE COMMUNITIES

KATIE E. SLOCOMBE*1, NICOLE LAHIFF1, CLAUDIA WILKE1 and CAT HOBAITER2

*Corresponding Author: katie.slocombe@york.ac.uk ¹Department of Psychology, University of York, York, UK ²School of Psychology and Neuroscience, University of St Andrews, UK

Referential gestures such as pointing are ubiquitous in human societies and the production and comprehension of referential gestures requires many of the cognitive capacities also crucial for language, including reading intentions and establishing common ground (Tomasello 2008). Given the importance of referential gesturing in human communication, there has been significant debate over the evolutionary origins of this behavior. Although captive and enculturated chimpanzees spontaneously learn to point referentially for human caregivers (Leavens et al., 2004), habitual use of referential gestures was thought to be absent in wild chimpanzees, until Pika and Mitani (2006) reported the use of 'directed scratches' during grooming bouts amongst male dyads in the Ngogo community of chimpanzees. In this community, after the recipient of grooming scratched his own body, in 64% of cases, the groomer then groomed the scratched location. Pika and Mitani concluded that wild chimpanzees use gestures to specify where they want to be groomed and that these scratch gestures were both referential and iconic. In an attempt to replicate and extend these findings we recorded dyadic grooming bouts in the Kanyawara and Sonso communities of chimpanzees, located approximately 10 and 200km from the Ngogo community in Uganda.

Grooming interactions were filmed and we coded all instances of 'big loud scratches' and 'present body part' gestures (where a body part is moved towards the groomer possibly to request grooming in that broad body area), of which only the former has been argued to be referential. We identified all scratches given by recipients of grooming and coded whether the groomer groomed within 5cm of the scratched location in the 10 seconds after the scratche ended. In cases where the groomer did not begin grooming in the scratched location, we then examined

whether the signaler repeated the scratch gesture in the same location (persistence) or presented the same body part to the groomer (elaboration) in the 30 seconds after the first scratch ended. Such persistence and elaboration behaviours would have indicated that the signaler's goal was to request grooming in a specific location.

We found little evidence of chimpanzees in Sonso and Kanyawara communities using 'big loud scratch' gestures to request grooming in a specific location. Very few scratches elicited grooming in the scratched location, and when the groomer failed to groom the scratched location, very few instances of persistence or elaboration to reach the goal were shown. This indicates that in these two communities of wild chimpanzees the signaler's goal is not to elicit grooming in a specific location. In contrast, we found 'present groom' gestures were highly successful in eliciting grooming on the presented body part and that big loud scratches were frequently used to initiate grooming bouts.

These findings show that the habitual use of referential gestures with conspecifics is not a universal aspect of chimpanzee communication and it highlights the importance of replicating important findings in multiple populations. Our findings raise the intriguing possibility that the referential use of the big loud scratch gesture is culturally specific to the Ngogo community, and thus more detailed analysis of this behavior and examination of the signaler's goals and circumstances that have given rise to this form of communication are required in this community. For now, it seems that when not faced with the 'referential problem space' that captivity brings (Leavens et al., 2005), not all wild chimpanzees habitually produce referential gestures and referential gesturing is not a ubiquitous feature of chimpanzee communication.

- Leavens, D. A., Hopkins, W. D., & Thomas, R. K. (2004). Referential communication by chimpanzees (Pan troglodytes). *Journal of Comparative Psychology*, 118(1), 48.
- Leavens, D. A., Hopkins, W. D., & Bard, K. A. (2005). Understanding the point of chimpanzee pointing: epigenesis and ecological validity. *Current Directions in Psychological Science*, 14(4), 185-189.
- Pika, S., & Mitani, J. (2006). Referential gestural communication in wild chimpanzees (Pan troglodytes). *Current Biology*, *16*(6), R191-R192.
- Tomasello, M. (2008). *Origins of human communication*. Cambridge, MA: MIT press.

DOES LEARNING FAVOUR COMMUNICATIVE EFFICIENCY?

KENNY SMITH*1 and JENNIFER CULBERTSON1

*Corresponding Author: kenny.smith@ed.ac.uk ¹Centre for Language Evolution, University of Edinburgh, UK

Natural languages are well-designed for communication, exhibiting structural properties which optimally trade off communicative function and efficiency (e.g. Zipf, 1949; Piantadosi et al., 2012; Jäger, 2007). These properties of language are typically explained as a consequence of modifications made by speakers during language use; in these accounts, learning plays no role in structuring languages to be communicatively optimal, and indeed simplicity biases in learning can erode communicative utility if unchecked (Kirby et al., 2008; Silvey et al., 2015).

However, several recent papers purport to show an improvement in communicative function through learning alone (e.g. Fedzechkina et al., 2012; Carstensen et al., 2015). Fedzechkina et al. (2012) show that participants trained on an artificial language featuring variable case marking will restructure that language in ways which would increase its communicative function, despite never using the language to communicate. They train participants over 4 days on an artificial language for describing events in which animate agents perform actions on animate or inanimate patients. The training language has variable word order, meaning that utterances are potentially ambiguous if the patient of an event is animate (but not if it is inanimate, since inanimates are never agents in their stimuli). The training language features case marking which, where it occurs, serves to eliminate this ambiguity: 60% of objects are case-marked, but animates and inanimates are equally likely to be case marked, yielding a distribution of case markers which is not communicatively optimal. Fedzechkina et al. found that learners increased the frequency of case marking on animate patients, and reduced it on inanimate patients. This results in a more communicatively efficient system, suggesting that learning may in fact play a critical role in structuring languages to be communicatively optimal, a result which runs contrary to the accounts reviewed above.

In Experiment 1 we replicated Fedzechkina et al.'s experiment and result (see Figure 1): we found that participants preferentially marked animate patients at later days. However, this replication highlighted a surprising feature of their stimuli: animates which are agents are never patients, and vice versa. By day 4 our participants were therefore able to correctly interpret sentences involving two animates with near-perfect accuracy even if the object was not case-marked; nonethe-



Figure 1. Proportion of case-marked objects in participants' productions in both experiments across days (testing occurred on days 2–4 only). The dashed line indicates frequency of case marking in the input, for both animate and inanimate patients; error bars indicate 95% CIs on the mean; individual points indicate by-participant means. Note the developing tendency to case-mark animate objects more often than inanimate objects, present in both experiments (as indicated by a significant day x animacy interaction in Experiment 1, and no interaction featuring Experiment across the full data set).

less they exhibited the pattern of case-marking that is putatively driven by the ambiguity of unmarked animate patients. We re-ran the experiment with a modified set of stimuli where all animates served as agents *and* patients, which should increase the utility of differential case marking. While Experiment 2 showed a modest increase in case marking on all objects, there was not a significantly stronger tendency to casemark animates (see Figure 1). Experiment 2 also featured an additional test on Day 4 where participants used the language to communicate with their alien language tutor; here, where communicative utility matters, we again saw an overall increase in case marking, but not preferentially on animates.

Overall, our results cast some doubt on the Fedzechkina et al.'s claim that biases in learning favour communicative utility: while their result is robust, this interpretation is at odds with results elsewhere in the literature, and makes predictions which do not appear to be borne out in their paradigm. We are currently exploring whether restructuring of the input in this case might be due to biases in learning which favour conditioning of variation (cf. Hudson Kam & Newport, 2009): their results may represent an artifactual case where biases in learning yield languages which are coincidentally better for communication. Thus, even if learning does not directly drive communicative function, it nevertheless plays an important explanatory role in accounting for the design of natural languages.

Acknowledgements

This project has received funding from the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme (grant agreement 681942).

- Carstensen, A., Xu, J., Smith, C. T., & Regier, T. (2015). Language evolution in the lab tends toward informative communication. in . In *Proceedings of the 37th annual meeting of the cognitive science society*. Austin, TX: Cognitive Science Society.
- Fedzechkina, M., Jaeger, T. F., & Newport, E. L. (2012). Language learners restructure their input to facilitate efficient communication. *Proceedings of the National Academy of Sciences, USA, 109*, 17897–17902.
- Hudson Kam, C. L., & Newport, E. L. (2009). Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology*, 59, 30-66.
- Jäger, G. (2007). Evolutionary game theory and typology. a case study. *Language*, 83, 74–109.
- Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: an experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences, USA*, 105, 10681-10686.
- Piantadosi, S. T., Tily, H., & Gibson, E. (2012). The communicative function of ambiguity in language. *Cognition*, 122, 280-291.
- Silvey, C., Kirby, S., & Smith, K. (2015). Word meanings evolve to selectively preserve distinctions on salient dimensions. *Cognitive Science*, *39*, 212-226.
- Zipf, G. K. (1949). *Human behaviour and the principle of least effort : an introduction to human ecology*. Cambridge, MA: Addison-Wesley.

UNDERSTANDING VS. DESCRIBING OTHERS' MINDS: MENTAL STATE LANGUAGE IN A SMALL-SCALE SOCIETY

ANDREW MARCUS SMITH^{*1}, ERICA A. CARTMILL¹, and H. CLARK BARRETT¹

*Corresponding Author: andrewmsmith@ucla.edu ¹Department of Anthropology, University of California, Los Angeles, USA

Theory of mind, or mentalizing, has been proposed to be a core cognitive domain underlying much of human social behavior (Corballis & Lea, 1999). Moreover, language functions as the primary channel through which social behavior is mediated (Carpenter et al., 1998; Clark, 1996; Tylen et al., 2010). As such, a complete understanding of either relies on an analysis of their interaction.

As a core cognitive domain, it is plausible that mentalizing has been subject to tight constraints imposed by natural selection. In line with this assumption, research on adult mentalizing has found relative homogeneity in neurotypical adult populations (e.g., Castelli et al., 2000; Senju et al., 2009). Such findings have bolstered claims suggesting there are few, if any, meaningful differences in adult mentalizing ability across both individuals and cultures. However, these same findings often come from primarily Western samples (see Liu et al., 2008 and Heyes & Frith, 2014 for exceptions), employ tests of mentalizing that show ceiling effects in adults (although see Turner & Felisberti, 2017 for a review of methods appropriate for adults) and equivocate competence and performance (cf. Wu & Keysar, 2007 for an exception). While human beings may have an implicit and species-typical capacity, or competence, to impute others' mental states, the way in which such imputations structure social behavior may differ across populations in accordance with prevailing socioecological conditions.

An emerging anthropological literature lends credence to this claim. Ethnographic accounts suggest that many small-scale societies have social norms that restrict talking about others' mental states (Robbins & Rumsey, 2008). Given the documented importance of mental-state talk for children's early sociocognitve development (Ruffman, Slade, & Crowe, 2002), these accounts, if true, raise a number of questions central to the evolution of language. Does the way in which

people talk about mental states actually vary across cultures and individuals? How does this variation relate to a presumably universal competence? To what extent does making knowledge of others' mental states explicit facilitate cooperation and communication? Despite the centrality of these questions, few quantitative studies of mental-state talk in a society with such norms have been conducted. Thus, we measured mental state language in a small-scale society with implicit norms against attributing mental states to others. An explicit quantification of mentalstate talk provides a metric against which ethnographic accounts can be measured. Should the data suggest differences in patterns of mental-state talk, subsequent studies are warranted. However, it is critical to first assess the truth of the claims.

A simple response-elicitation task was administered in Spanish to bilingual Shuar / Spanish speakers in a small-scale, hunter-horticulturalist society in Amazonian Ecuador (N=40, 20 female) as well as in English to a sample of American undergraduates (N=20, 15 female). Participants were shown a set of eight silent videos. Each set contained six videos designed to elicit mental state language (MS) and two videos that served as controls (C). MS videos depicted naturalistic interactions between multiple agents and featured richly mentalistic content, such as one character tricking another character. C Videos depicted causal, but inanimate agents interacting. Each video was 30-45 seconds long.

After each video, participants described what happened in the scene. Participants received a minimal prompt to reduce task demands and capture variation in the extent to which mental states were reported. Shuar participants were expected to describe scenes using less mentalizing language than American participants. Verbal responses were transcribed and coded according to a scheme adapted from Castelli et al. (2000) and Ruffman et al. (2002). Word counts for affective states, desire, epistemic states, and perception were obtained. Counts per description were scaled by description length to control for differences in verbosity. Hierarchical Poisson Regression models of word counts were run with culture and video type as fixed factors and participant as a random factor.

American participants were found to use words attributing perception, affective states, and epistemic states to characters significantly more often than were Shuar participants, while Shuar participants used desire words more frequently. These data confirm that our Shuar participants were less likely than Western participants to describe scenes they had witnessed in terms of an agent's mental states. These data are important for a complete understanding of the evolution of language. Language cannot be understood as apart from the mentalizing capacity from which it follows – the two are inextricably linked and undergird human beings' uniquely elaborated sociality.

- Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. Monographs of the society for research in child development, i-174.
- Castelli, F., Happé, F., Frith, U., & Frith, C. (2000). Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. Neuroimage, 12(3), 314-325.
- Clark, H. H. (1996). Using language. Cambridge university press.
- Corballis, M. C., & Lea, S. E. (1999). The descent of mind: Psychological perspectives on hominid evolution. Oxford University Press
- Heyes, C. M., & Frith, C. D. (2014). The cultural evolution of mind reading. Science, 344(6190), 1243091.
- Robbins, J., & Rumsey, A. (2008). Introduction: Cultural and linguistic anthropology and the opacity of other minds. Anthropological Quarterly, 81(2), 407-420
- Ruffman, T., Slade, L., & Crowe, E. (2002). The relation between children's and mothers' mental state language and theory-of-mind understanding. Child development, 73(3), 734-751.
- Senju, A., Southgate, V., White, S., & Frith, U. (2009). Mindblind eyes: an absence of spontaneous theory of mind in Asperger syndrome. Science, 325(5942), 883-885.
- Turner, R., & Felisberti, F. M. (2017). Measuring Mindreading: A review of behavioral approaches to testing cognitive and affective mental state attribution in neurologically typical adults. Frontiers in psychology, 8.
- Tylén, K., Weed, E., Wallentin, M., Roepstorff, A., & Frith, C. D. (2010). Language as a tool for interacting minds. Mind & Language, 25(1), 3-29.
- Wu, S., & Keysar, B. (2007). The effect of culture on perspective taking. Psychological science, 18(7), 600-606.

HOW SOCIOLINGUISTIC THEORY CAN INFORM CULTURAL EVOLUTION, AND HOW TO TEST IT

BETSY SNELLER¹ and GARETH ROBERTS*1

*Corresponding Author: gareth.roberts@ling.upenn.edu ¹Department of Linguistics, University of Pennsylvania, Philadelphia, PA, USA

In cultural evolution, behavioral variants can be governed by neutral forces or selection, in which individuals are biased in some way. An important bias derives from social meaning: People may select variants that *index* socially desirable traits. Silverstein (2003) and Eckert (2008) distinguished between different *orders of indexicality* that a cultural variant may be associated with. First-order indices identify a speaker as belonging to a particular group (defined, e.g., by region or class). A shift to second-order indexicality occurs when variants come to index perceived traits of a group, such as toughness. Second-order indices differ from first-order indices along two dimensions: First, they are more *alienable*, less inherently connected to a particular group of people. Second, traits indexed by second-order variants are not arbitrary; they are *socially relevant* characteristics set up in opposition to perceived traits of other groups.

This has consequences for cultural evolution. All things being equal, firstorder indices should spread neutrally; second-order indices should be selected for, propagating at a greater rate. We tested this using an artificial language paradigm devised by Roberts (2010), originally to investigate new-dialect formation. Groups of four participants (160 in total) sat in separate cubicles with computers and played a game in which they took on the role of one of two species of alien: the weaker Wiwos or the (explicitly) tougher Burls. The two species were taught different dialects of the same artificial language (which consisted of twelve words with two or three CV syllables; one dialect would have b where the other had f). The game consisted of a series of rounds in each of which players were paired up, their species made clear to each other. At the start of the round players chatted in the alien language using text-based instant messaging. Then they could choose to trade resources with each other or (in certain conditions) challenge each other to fights. Each player started with 22 points' worth of resources; the goal was to have the most resources at the end of the game. This could be achieved by trading (where anything given was worth double to the receiver) or, in relevant conditions, by challenging one's partner to a fight. The partner could agree to the fight or run away. If a player lost or ran away, they would lose resources and the

other player would gain some. Running away was less bad than losing a fight, and scaring an opponent more profitable than winning a fight.

We manipulated the alienability and social relevance of a consonantal feature of the Burl dialect in a 2×2 design. We manipulated alienability by giving Wiwos an explicit stereotype about the Burl dialect. In inalienable conditions they were told "Burls tend to use *b* instead of *f*". In alienable conditions they were told "Tougher aliens tend to use *b* instead of *f*." We manipulated social relevance by including or not including fighting in the game. See Table 1 for a summary with condition names.

Table 1. Conditions of the experiment

	"Tougher aliens"	"Burls"
Fighting	Second-order	Relevance-only
No fighting	Alienability-only	First-order

We then measured the rate at which Wiwos used the Burl consonant. We predicted they should do so more in the Second-order condition (where it was associated with an alienable and socially relevant trait, toughness) than in the first-order condition. As can be seen in Figure 1, the Burl consonant was borrowed to a significantly greater extent in the Second-order than the First-order condition. Interestingly, the other two conditions behaved just like the First-order condition. This study thus provides an empirical test of theoretical notions devised in sociolinguistics and has important consequences for our understanding of how social meaning influences the propagation of cultural variants.



Figure 1. Proportion of Burl consonants used by Wiwos in each condition

- Eckert, P. (2008). Variation and the indexical field. *Journal of Sociolinguistics*, 12(4), 453–476.
- Roberts, G. (2010). An experimental study of the role of social selection and frequency of interaction in linguistic diversity. *Interaction Studies*, *11*(10), 138–159.
- Silverstein, M. (2003). Indexical order and the dialectics of sociolinguistic life. *Language and Communication*, 23, 193–229.

"THE GREATEST STORY EVER TOLD". SEMIOSIS EMERGING FROM MIMESIS AND/OR NARRATIVITY

GÖRAN SONESSON^{*1,}

*Corresponding author: goran.sonesson@semiotik.lu.se ¹Cognitive Semiotics, SOL, Lund University, Lund, Sweden

Language may emerge as one of several specializations out of mimesis, understood as whole-body communication. Starting out from the idea of a mimetic stage as the precursor to language as first posited by Donald, and building on the observations on this notion made by Zlatev and Sonesson, we consider two recent proposals for rendering mimesis more specific. The first of these proposals, due to Żywiczyński, & Wacewicz, aims to narrow down mimesis to pantomime, of which they try to give a precise definition. While this proposal throws much light on what is involved, pantomime in this sense can, in the end, hardly be distinguished from (communicative) mimesis, given an appropriate definition. The other proposal, due to Ferretti et alia, suggests that narrative thinking is at the origin of both mimesis and language. We argue that, while narrativity builds on the specific character of human temporal cognition, the latter cannot become narrative without first being manifested by a semiotic resource.

1. Setting the Scene

According to the evolutionary theory propounded by Donald (1991, 2001), *episodic memory*, the memory for single situated happenings, is something that human beings share at least with apes. *Mimetic memory*, or perhaps rather the peculiarly human form of mimetic memory, is restricted to human beings and their ancestor species. Donald terms the third stage *mythic*, because it involves the construction of narratives, possibly used to recount myths, this capacity being, in his view, at least one of the reasons why language evolved. Perhaps this third stage could be understood to be intermediate between biological and cultural evolution, but the fourth stage, termed *theoretical*, by Donald, is where evolution breaks free of biology, producing pictures, writing, and theorizing. Donald (2010) places the latter semiotic resources at this stage because they are "exograms", i.e., memory records which have an existence independently of any (specific and current) subject. There are excellent reasons to think that narrativity predates language, at least in a rudimentary form, being manifested

both in gesture and in (not endurable forms of) pictures (See Sonesson (2007, 2016). Since then, there have been proposals, on the one hand, to consider pantomime (which, in Donald's scheme, is only one of the manifestations of the mimetic stage) to be at the origin of the process leading to the emergence of language (Żywiczyński, & Wacewicz, 2017; & Sibierska 2016), and, on the other hand, to suggest that narrativity as such is at the origin of language (Ferretti, Adornetti, & Chiera, 2017). In the next section, we will discuss the first proposition, reserving the discussion of the latter for the third section.

2. Miming Mimesis

Donald (1991: 170ff) lists a number of criteria, which have to be fulfilled by any item pertaining to be an instance of mimesis: intentionality (in the sense of having a purpose), generativity (in the sense of being divisible into components and recombined), communicability (which Donald goes on to explain as public availability, which is really a more limited notion), referentiality (in the sense that "the referential act must be distinguished from its referent"), potentially involving an unlimited number of objects and events, and autocuing (i.e. being volitionally controlled). All through his papers, Donald (e.g. 1991, 2001, 2010) exemplifies this stage using terms current in the vernacular, some of which are rather vague and ambiguous, such as imitation, pantomime, and gesture, and others which seem to have only partly overlapping meanings, such as tool use and skill.

Sonesson (2007, 2016) pointed out that bodily movements that are referential, that is, which form signs, must be distinguished from action which impinges on the material world, such as tool use and skill, even though the latter may be propagated by means of imitation. For the definition of the sign, Sonesson relies on notions formulated by Husserl and Piaget: 1) the sign contains (a least) two parts (expression and content) being as a whole relatively independent of that for which it stands (the referent); 2) these parts are differentiated, from the point of view of the subjects involved in the semiotic process, even though they may not be so objectively, i.e. in the common sense *Lifeworld* (except as signs forming part of that Lifeworld); there is a double asymmetry between the parts, because one part, expression, is more directly experienced than the other; and because the other part, content, is more in focus than the other. None of this applies to tool use and skill.

Zlatev (2007, 2014) proposed a division of the notion of bodily mimesis, distinguishing between "dyadic mimesis" which includes imitation and "triadic mimesis" which brings along communicative signs. Zlatev, Donald, and

Sonesson (2010) together proposed that bodily mimesis, involving performances of the whole body, must have been at the origin of semiotic evolution, which seems to have made human beings into a very special kind of animal. The experiment realised by Zlatev, Wacewicz, Zywiczynski, & van de Weijer (2017), in which action employing both sound and gesture turned out to be less efficacious in communicating specific content than gesture alone, can be taken to suggest that multiple sense modalities do not constitute an ideal point of origin, but it is still possible that human semiosis began as action tailored to the body as a whole.

Whatever the fate of their "pantomime-first" proposal, we owe to Zywiczynski, Wacewicz, & Sibierska (2016) an attempt to define pantomime, as something more specific than Donald's mimetic stage and Zlatev's bodily mimesis. Pantomime, in their sense, is mimetic, (iconically) motivated and nonconventional, improvised, (primarily) visual, but potentially multimodal, holistic and involving the use of the whole body. In order to be a precursor to language, pantomime must also, they claim, be communicatively complex, self-sufficient and semantically advanced, in the sense of being displaced, open-ended and semantically universal. If mimesis is here taken in the sense of Donald, it would seem that this already contains some of the properties listed: open-endedness appears to correspond to the reference to unlimited domains of objects, generativity, and perhaps also semantic universality. In fact, the term generativity certainly appears to go in this sense, but as Donald defines it, it seems to be in contradiction to the holistic character postulated by our authors. Surprisingly, "motivated and non-conventional" is not part of Donald's criteria, although it is implied both by the term and the examples given. It is true, however, that some of the capacities Donald qualify as mimetic, such as tool use and skill, are not iconic, except in plausibly being acquired through imitation.

Iconicity may not be as problematic as some authorities within philosophy and semiotics have claimed (see Sonesson 1989). On the other hand, nonconventionality certainly is, if it is taken to imply "some level of universality as opposed to culture-specificity" (Zywiczynski, Wacewicz, & Sibierska 2016, no page number). We can obviously not contradict this claim by producing any examples of pantomime, since the latter is supposed to be improvised. But some iconic signs in signed languages may well have an origin in such pantomime, without being necessarily universal. In the notorious case of the sign for "tree", ASL uses the lower arm to imitate a stem with branches (apparently caught in a hurricane), and Danish sign language employ both hands to outline the contours of the treetop and the trunk, while Chinese sign language have the hands outlining the trunk from the bottom to the top (See Klima & Bellugi 1979: 21f). All these signs are motivated, although the rely on different properties of the object referred to. A more general point may be that, to the extent that praxis for everyday behaviour is different in different societies (e.g. different ways of eating, sitting, greeting, and perhaps even perceiving the world, etc.), the corresponding iconical rendering will be different.¹

Pantomime certainly is improvised in the sense of being autocued, but something more is implied by the term improvisation: that the signs are all the time created anew. If it is true that there are "languageless adults in Mexico – deaf, never taught sign language, living together on the fringes of society – /who/ mime narratives for one another" (Boyd 2009: 130f; quoted in Ferretti, Adornetti, & Chiera, 2017), it seems probable that in such a context pantomime will rapidly be standardized, and indeed give birth to a signed language, making the claims of "languageless" deaf people highly suspect. Indeed, turn-taking, said by Żywiczyński, & Wacewicz (2017) to be instrumental in turning pantomime into verbal language, would rather rapidly erode the spontaneity of such an improvisation. In fact, the idea of pantomime being improvised conveys up images of the occasional meetings of Neanderthals and *Homo sapiens*, likened by Fagan (2011: 16ff) to the meeting between the TV team led by David Attenborough and the San people, although Fagan only allows for "smiles and eyebrows that could frown, signify wonder or disapproval, or ask a question".

Another problematic criterion is that of whole-body involvement, which may extend to the environment (see Żywiczyński, Wacewicz, & Sibierska, 2016). As the authors admit, however, "movements of the hand and arm do have a rather critical part to play in pantomime, but only to the extent that they are implicated in the holistic generation of meaning" – or, as they also frame it, as being part of a system (Żywiczyński, & Wacewicz 2017). If we understand the term system here in the structuralist sense, it can clearly not involve the paradigmatic axis, since pantomime is supposedly improvised, but it would have to take place on the syntagmatic axis. Whatever the syntagmatic connection amounts to, nevertheless, it cannot be syntax. That leaves us wondering what kind of connection it is.

More could be said about, for instance, the problematic criterial nature of holism, but, for the time being, let us only salute the inclusion of the criterion of displacement, well-known from Hockett's design features of language, and more

¹ If the term sconventionality is here only meant to suggest that there are right and wrong ways of realising the pantomimic sign, then normativity may be a more adequate term.

generally, from Piaget's notion of cognition. This means that pantomime will not include "interactions between chimpanzees and zoo visitors" in which "the two species imitated each other" about 10% of the time (see Persson, Sauciuc, Alenkær Madsen, 2017). Still, in the end it seems that Żywiczyński, Wacewicz, & Sibierska (2016) have only succeeded in "unpacking", as they themselves put it, the notion of mimesis, at the price of making it more vulnerable to criticism – which, in the game played by scientists, is not a small contribution. Perhaps, in the end, we should retain the term pantomime for that which is mimesis not only in the sense of Donald and Zlatev, but in the sense of the ancient Greeks (See Sörbom 1966). Aristotle, for instance, would claim, according to Ranta (2000: 68) that

"X is a mimetic object if X represents, and is (to some extent) similar to (mentally imagined) types of perceivable or imagined objects, subjects, or actions."

To the extent that this definition is applied to all kinds of art, not just plastic art, it must involve some very abstract kind of similarity indeed. In this sense, Auerbach (1946) was not wrong in using the term mimesis to describe literary realism. With reference to Auerbach, however, Donald (1991: 170) points out that Auerbach's subject is not "purely mimetic".

Another question is what might bring us from pantomime to language. As we suggested above, the introduction of turn-taking must have occurred rather rapidly, and cannot be sufficient to account for the passage from mimesis to language. Another factor, mentioned by Żywiczyński, & Wacewicz (2017) is the "platform of trust". This is certainly an excellent gauge for spelling out the difference between human beings and other animals: many experiments with primates have shown that apes, contrary to human children, do not understand the use of pointing and other semiotic vehicles, probably because they cannot imagine that anybody would help them find the food they desire (e.g. Zlatev et al. 2013 and literature referred to there). Nevertheless, it would seem that trust is a much more general factor distinguishing human beings from animals (although some trust must certainly exist among all animals living in groups). Indeed, trust was used to define the nature of society by Garfinkel (1963), the founder of ethnomethodology, who took his inspiration from the phenomenologists Schütz and Gurwitsch.²

² From a Darwinean point of view, it seems that only group selection can account for the survival value of trust, whether or not you postulate a cheater module. Thus, it would define Ego-culture, in the sense of Sonesson (2012; 2016a, b). But that is not our present concern.

3. Manually Retelling the Tale

Perhaps narrativity, then, is the key to the passage from mimesis to language. It happens to be the other part of the definition of art (which we must take as equivalent to semiotic resources) propounded, in different ways, by Plato and Aristotle, i.e. diegesis (See Sörbom 1966). After all, from Lessing to Donald, all authorities have claimed that only language can tell a story, or at least that language does it best. Lessing, who is more explicit about his argument, says that language, since it consists of signs in time, is better at rendering temporal facts, while pictures, which are signs in space, may have an advantage as far as rendering spatial facts is concerned. There are problems with this description, which do not have to concern us here (see Sonesson 1997; 2014). Mimesis, gesture, pantomime, and anything of the kind are immune to this critique, if it is at all relevant, because, to the extent that they are signs, they certainly are temporal at the level of expression, which is what counts here.

According to Ferretti et alia (2016a, b; cf. Ferretti & Adornetti 2016), narrativity is not only manifested already in mimesis, but, before that, there is narrative thinking. A lot could be said about this claim relative to the criteria distinguishing narrativity and narrativehood, as characterized by Prince, Ryan, and others (see Sonesson 1997 and below), but, at present we are concerned with a much more fundamental distinction. Most narratologists take it for granted that a narrative is built up from some kind of semiotic resource, normally language, which represents some other level of reality, however fictive. In the so-called Paris school, inspired by the work of Greimas, narrativity, nevertheless, is applied also to events taking place in the real world of our experience. Curiously, the abundant use of the term "representation" in cognitive science seems to lead to the same confusion.

If we follow classical narratology in defining narrative, not as a sequence of events, but as an "external" representation of such a sequence, we still have to face the rejoinder that the single subject may be telling himself a tale. The question is what this means. As observed by Sonesson (2015), animals like the tick only need to be aware of reality in terms of before and after (McTaggart's B-series). Mammals generally, and at least some birds, clearly live in a world in which time is counted in terms of past, present and future (McTaggart's A-series), which requires the insertion of the ego into the stream of consciousness. Moreover, some animals are able to distance themselves from the streaming of the stream of consciousness, initiating themselves (i.e. autocuing) the acts of remembrance and anticipation, or, in other terms, "mental time travel". Indeed, Ferretti et alia (2016a, b) refer to "mental time travel", to justify the idea of
narrative thinking preceding any other manifestation of narrativity.

While the ability to accomplish acts of remembrance and anticipation certainly requires an advanced ability to distance oneself from one's own stream of consciousness (once one possesses one), a further level of distancing would probably be required in order to collegiate the acts of remembrance and anticipation in the form of a narrative sequence, that is, to form a minimal story, as defined by Prince (1982: 1): "at least two events with a temporal link on the content side" - which, as observed by Sonesson (1997), presupposes there to be an expression side, which does not, however, necessarily have a temporal link. The idea of a temporal link could be explicated in the following way: given two state descriptions which pertain in some sense to the same piece of the world, the temporal moment described in one of them is different from that in the other. If there is such a link in mimesis, however, it would necessarily have to be parallel in expression and content, while this is (*pace* Lessing) rarely the case in language (See Sonesson 2014). Thus, language is able to externalize the mental acts of remembrance and anticipation, in the shape of flashbacks and flashforwards. Donald's claim that narrative is wedded to language in the mythic stage can be understood in the sense that each stage is characterized by the properties it brings to full fruition. On the other hand, myths may not be prominent for containing flashbacks and flash-forwards - although they are endemic already in Homer.

Sonesson (1997) has argued that some of the criteria which, according to Prince (1996), bring us from mere narrativehood to narrativity, can be realised by static pictures, while others cannot. It may be difficult for a single picture to show logically unpredictable antecedents or consequences; deep causality (first and last events linked in significant ways); elements of conflict between different subjects; etc. Pantomime would seem to be closer to language in these respects. Like language and pictures, pantomime may be able to illustrate transactiveness (actions as opposed to happenings) and transitiveness (events involving agent and patient). However, because of its rough way of creating similarity, it may be worse off than pictures for rendering specificity instead of generality (the opposite of sequences fitting any or indefinitely many sets of circumstances) and the presence of virtual actions (what could have happened but did not — alternative courses of action).³

It is, of course, also possible for cognition to fulfil these criteria. But if the result of these cognitions remains in thought only, there is no expression

³ Experimental evidence would need to be adduced to enrich our understanding of these issues.

corresponding to this content. Such a claim, naturally, will have to rely on Husserl's (1980) argument, taken up again with empirical references, by Thomson (2007), that mental images are not really images, but reactivations of the corresponding actions.

4. Conclusion

As a preliminary conclusion, it might be better to think of language and all other kinds of semiosis emerging from mimesis in the general sense of whole-body communication (Zlatev, Donald, and Sonesson 2010). In this sense, mimesis contains the embryo not only of pantomime, language, and narrativity, but also of depiction, each of which eventually become specialized into their own semiotic resources. In fact, when Arbib (2012: 219) says that pantomime has its limitations, because it is hard to pantomime "blue", this is hardly an argument for the development of language. You can invent a gestural emblem for blue, just as easily as you can coin a word for it, but in none of these cases will you really be able to convey the experience of blueness, if it is not known beforehand. Indeed, as Lessing already said, only depiction uses signs the expressions of which are shapes and colours in space (Sonesson 2014). One could imagine drawing emerging in evolution out of the traces left on a surface by gestures, as is the case in child development. This makes sense, since we know that, in many cultures, sand painting is almost as transient an event as gesture. Pictures as exograms, in Donald's fourth stage, is a later outgrowth of depiction. But it may very well be the first instance of this evolvement.

If so mimesis, rather than pantomime or narrativity, will really be at the origin of it all. Thus, what is from the beginning a rather undifferentiated wholebody communication later divides up according to what each semiotic resource does best: language for telling stories, depiction for showing what blue and square look like, and gesture perhaps mostly as a complement to speaking and, more exceptionally, as a substitute for it, in the shape of signed language or pantomime.

Acknowledgements

The author wants to thank Jordan Zlatev and two anonymous reviewers for comments on an earlier draft of this paper.

- Arbib, M. (2012). *How the brain got language*. Oxford University Press, Oxford.
- Auerbach, E. (1946). *Mimesis: dargestellte Wirklichkeit in der abendländischen Literatur*. Bern: Francke.
- Boyd, B. (2009). On the origin of stories: evolution, cognition, and fiction. Cambridge, Mass.: Belknap Press of Harvard University Press.
- Donald, M. (1991). Origins of the modern mind: Three stages in the evolution of human culture. Cambridge, MA: Harvard University Press.
- Donald, M. (2001). *A mind so rare: The evolution of human consciousness*. New York: Norton.
- Donald, M. (2010). The exographic revolution: Neuropsychological sequelae. In L. Malafouris & C. Renfrew (Eds.), *The cognitive life of things: Recasting the boundaries of the mind* (pp. 71-79). Cambridge: McDonald Institute Monographs.
- Fagan, B.M. (2011). Cro-Magnon: how the Ice Age gave birth to the first modern humans. (Paperback ed.) New York: Bloomsbury Press.
- Ferretti, F. & Adornetti, I. (2016) Mindreading, Mind-travelling, and the Proto-Discursive origins of language. In Zlatev, J., Sonesson, G., & Konderak, P. (eds.), *Meaning, Mind and Communication. Explorations in Cognitive Semiotics.* Frankfurt/M.: Peter Lang.
- Ferretti, F. Adornetti, I., Chiera, A., Nicchiarelli, S., Magni R., Valeri G., Marini, A. (2017a) Mental Time Travel and language evolution: a narrative account of the origins of human communication. In *Language Sciences* 63 (2017) 105–118
- Ferretti, F., Adornetti, I. & Chiera, A. (2017b). The Narrative Dimension Of Pantomime: An Evolutionary Perspective. Lecture given at the minisymposion "The Role of Pantomime in Cognitive-Semiotic Evolution" at Lund University, September 8, 2017.
- Garfinkel, H. (1963) A Conception of, and Experiments with, "Trust" as a Condition for Stable Concerted Actions, in O.J. Harvey (ed.) *Motivation and Social Interaction* (pp. 187-238). New York: Ronald Press.
- Husserl, E. (1980). *Phantasie, Bildbewusstsein, Erinnerung*. Husserliana XXIII. The Hague: Nijhoff
- Klima, E.S., Bellugi, U. & Battison Robbin (1979). *The signs of language*. Cambridge, Mass.: Harvard University Press.
- Persson, T., Sauciuc, G., Alenkær Madsen, E. (2017). Spontaneous crossspecies imitation in interactions between chimpanzees and zoo visitors. *Cross-Mark: Primates*. DOI 10.1007/s10329-017-0624-9
- Prince, G. (1982). *Narratology: the form and functioning of narrative*. Berlin: Mouton.
- Prince, G. (1996). Remarks on Narrativity in C. Wahlin, (ed.) Perspectives on Narratology (pp 95-106). Frankfurt: Peter Lang.

Ranta, M. (2000). *Mimesis as the representation of types: the historical and psychological basis of an aesthetic idea*. Diss. Stockholm : Univ.. Stockholm.

Sonesson, G. (1989). Pictorial concepts. Lund: Lund University Press.

- Sonesson, G. (1997). Mute narratives: New issues in the study of pictorial texts. In U. B. Lagerroth, H. Lund & E. Hedling (Eds.), *Interart poetics: Essays on the interrelations of the arts and the media* (pp. 243-252). Amsterdam: Rodopi.
- Sonesson, G. (2007). From the meaning of embodiment to the embodiment of meaning. In T. Zimke, J. Zlatev & R. Frank (Eds.), *Body, language and mind. Vol 1: Embodiment* (pp. 85-128). Berlin: Mouton.
- Sonesson, G. (2012). Between homeworld and alienworld: A primer of cultural semiotics. In E. W. B. Hess-Lüttich (Ed.), Sign culture – Zeichen: Kultur Festschrift for Roland Posner (pp. 315-328). Würzburg: Königshausen & Neumann.
- Sonesson, G. (2014). Translation and Other Acts of Meaning. In Between Cognitive Semiotics and Semiotics of Culture. In *Cognitive Semiotics*, 7:2, 2014, 249-280.
- Sonesson, G. (2015). From remembering to memory by way of culture: A study in cognitive semiotics. *Southern Journal of Semiotics*, 5(1), 25–52.
- Sonesson, G. (2016a). Lifeworlds: The Cognitive Semiotics of Culture. In Dunér, D., & Sonesson, G. (eds.), *Human Lifeworlds: The Cognitive* Semiotics of Cultural Evolution (pp. 23-62). Frankfurt/M.: Peter Lang.
- Sonesson, G. (2016b). Cultural Evolution: Human History as the Continuation of Evolution by (Partially) Other Means. In Dunér, D., & Sonesson, G. (eds.), *Human Lifeworlds: The Cognitive Semiotics of Cultural Evolution* (pp. 301-336). Frankfurt/M.: Peter Lang
- Sörbom, G. (1966). *Mimesis and art: studies in the origin and early development of an aesthetic vocabulary*. Diss. Uppsala : Univ.. Stockholm.
- Thompson, E. (2007). *Mind in life: biology, phenomenology, and the sciences of mind*. Cambridge, Mass.: Belknap Press of Harvard University Press.
- Zlatev, J. (2007). Language, embodiment and mimesis. In T. Ziemke, J. Zlatev & R. Frank (Eds.), *Body, language, mind. Vol 1: Embodiment*, (pp. 297-338). Berlin: de Gruyter.
- Zlatev, J. (2014). Bodily mimesis and the transition to speech. In M. Pina & N. Gontier (Eds.), *The evolution of social communication in primates* (pp. 165-178). Berlin: Springer.
- Zlatev, J., Madsen, E. A., Lenninger, S., Persson, T., Sayehli, S., Sonesson, G., van de Weijer, J. (2013). Understanding communicative intentions and semiotic vehicles by children and chimpanzees. *Cognitive Development*, 312–329.
- Zlatev, J., Donald, M., and Sonesson, G. (2010). From body to mouth (and body). In A. Smith, M. Schouwstra, B. deBoer, and K. Smith, *The evolution*

of language (pp. 527-528). London: World Scientific.

- Zlatev, J., Wacewicz, S., Zywiczynski, P. & van de Weijer, J. (s2017) Multimodal-first or pantomime-first? Communicating events through pantomime with and without vocalization. Interaction Studies 18:3 (2017), 455–479.
- Żywiczyński, P., Wacewicz, S., & Sibierska, M. (2016) *Defining Pantomime for Language Evolution Research*. Consulted online on September 11, 2017: https://link.springer.com/article/10.1007/s11245-016-9425-9/fulltext.html
- Żywiczyński, P. & Wacewicz, S. (2017) *Pantomime and language evolution*. Lecture given at the minisymposion "The Role of Pantomime in Cognitive-Semiotic Evolution" at Lund University, September 8, 2017.

LANGUAGE COMPLEXITY AS AN INTERACTION BETWEEN SOCIAL STRUCTURE, INNOVATION, AND SIMPLICITY

MATTHEW SPIKE

matthew.spike@anu.edu.au Centre of Excellence for the Dynamics of Language, Australian National University

There are several claims regarding how social factors determine linguistic complexity: the world's most complex languages seem to be spoken by small, isolated populations, with dense social networks and a high proportion of child learners (see Trudgill, 2011), and simpler languages are spoken by larger populations (e.g. Lupyan & Dale, 2010), in language contact situations (e.g. Trudgill, 2011) and in particular emerge in the formation of creoles (e.g. McWhorter, 2001). Whatever the empirical status of these claims, the causal mechanisms offered up by these accounts are often conflicting. For example, is simplification driven mainly by speakers (e.g. Wray & Grace, 2007) or L2 learners (e.g. Trudgill, 2011)? Is complexification the normal directional trend in languages which are 'left alone' (Trudgill, 2011, p.235), or are complex features *selected for* by child learners (Lupyan & Dale, 2010)? Modelling work by Reali et al. (2014) suggests that simpler (i.e. more easily learned) conventions will predominate in larger populations, but how does this interact with differently structured populations?

I present results from a simple computational model which investigates the effect of all these factors, and argue that all these mechanisms can be understood in terms of their contribution towards i) maintaining complexity, and ii) the fixation of novel complexity in the language of a sub-population. This work supports some conclusions found in the established literature — for example, both speakers and L2 learners can drive simplification — but also provides a number of novel insights: for example the importance of *innovation* and its impact on different populations.

Individual agents are modelled as collections of tokens — a modified Moran model (Moran, 1962) or Pólya-Hoppe urn (Hoppe, 1984), where each token represents a specific *convention*, after Reali et al. (2014). However, where Reali et al. assume that complexity is an *inherent* property, capturing a difference between *easy* and *hard* conventions, the literature usually characterises complexity as *systemic* property, i.e. an increase in irregularity, redundancy, and the number of categories, along with a decrease in obvious systematicity (Trudgill, 2011). As such, complexity in this model is expressed quantitatively, i.e. the number of conventions.

tions which are shared across whole and sub-populations. This corresponds with the idea that the more complex the language, the more information is required to acquire it.

The parameters of the model allow us to compare the effects of varying i) the overall population size; ii) the population structure (complete, random, small-world, and connected-caveman graphs); iii) an *innovation* bias, (the probability of an agent originating a novel convention); iv) a *simplicity* bias which abstracts over production, learning, and memory (a preference for using majority over minority variants); and v) the L1/L2 ratio, (modelled in two ways: a) L1 agents having more linguistic exposures, and b) different degrees of heterogeneity in the initial states of the population).

There are several main results: first, the overall population size is less important than the structure of that population, which determines the effective population size for any given agent. Simply put, if the agent is connected to many others via relatively few steps (i.e. a low average shortest path length), then its effective population size is large, and vice versa. Second, whereas Trudgill predicts that smaller populations should maintain high levels of complexity, results show drift effects and a subsequent loss of complexity taking over unless there is a strong anti-simplicity bias. Larger effective populations are more able to maintain complexity unless speakers or learners have a strong preference for simplicity. Third, the innovation of novel complexity plays an important dual role: in small effective populations, novel complexity fixates easily, in line with the strength of the bias. In larger populations, on the other hand, fixation of novel complexity becomes impossible and, surprisingly, high levels of complexity innovation have the paradoxical effect of leading to a *collapse* in complexity unless the simplicity bias is strong, as this overwhelms the population-wide system with noise. Finally, highcontact situations such as found in the emergence of creole languages only lead to simplification in large effective populations with a strong bias towards simplicity: in other conditions, such as when populations are small or the simplicity bias is weak, then we can see an *increase* in complexity.

In general, the simplicity bias acts against both the maintenance of existing complexity and the fixation of novel complexity, while the innovation bias promotes that fixation but works against maintenance. These both interact with the effective population size: larger populations promote the maintenance of existing complexity, but prevent the fixation of novel complexity. These results suggest that we should look to populations which speak highly complex languages for evidence of individual biases for maintaining complex forms and innovating new forms. Finally, we should consider the possibility that these biases are themselves culturally inherited, and so vary between social and linguistic groups.

References

Hoppe, F. M. (1984). Pólya-like urns and the Ewens' sampling formula. Journal

of Mathematical Biology(20), 91–94.

- Lupyan, G., & Dale, R. (2010, jan). Language structure is partly determined by social structure. *PloS one*, *5*(1), e8559.
- McWhorter, J. H. (2001). The world's simplest grammars are creole grammars. *Linguistic Typology*, *5*, 125–166.
- Moran, P. A. P. (1962). *The statistical processes of evolutionary theory*. Clarendon Press.
- Reali, F., Chater, N., & Christiansen, M. H. (2014). The paradox of linguistic complexity and community size. In S. . Roberts, H. . Lyn, & H. . Cornish (Eds.), *Evolang 10* (pp. 270–277). Vienna: World Scientific.
- Trudgill, P. (2011). Sociolinguistic typology: Social determinants of linguistic complexity. Oxford University Press.
- Wray, A., & Grace, G. (2007). The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua*, 117(3), 543–578.

I SEE WHAT YOU DID THERE: THE ROLE OF ICONICITY IN THE ACQUISITION OF SIGNS

ASHA SATO*1, MICHAEL RAMSAMMY2, MARIEKE SCHOUWSTRA1 and SIMON KIRBY1

*Corresponding Author: Asha.S.Sato@ed.ac.uk ¹Centre for Language Evolution, University of Edinburgh, Edinburgh, UK ²Department of Linguistics and English Language, University of Edinburgh, Edinburgh, UK

Recent theoretical syntheses offer a view of language in which iconicity -aperceived resemblance between form and meaning - is seen as a fundamental design feature alongside arbitrariness (Dingemanse et. al. 2015). Under this view, iconicity serves to bootstrap acquisition, and there is a large body of work from both spoken and gestural modalities confirming that iconic signs are easier to acquire than arbitrary signs (for an overview, see Lockwood & Dingemanse, 2015; Perniss et. al. 2010). However, two recent studies suggest a more nuanced picture of iconicity's contribution to learning: In an artificial language learning experiment using a whistled language, Verhoef et. al. (2016) found that whistles were reproduced less accurately in a condition where iconicity was possible compared to a condition where iconicity was disrupted by scrambling the correspondence between signals and meanings. Similarly, in a longitudinal study of phonological development in British Sign Language (BSL) learners, Ortega & Morgan (2015) found that learners produce iconic signs with less articulatory accuracy than arbitrary signs of equal complexity. These two results are apparently contradictory to the idea that iconicity provides a learning advantage, but we suggest this is because most iconicity learning studies have focused on the acquisition of the *mapping* between form and meaning, thus potentially obscuring subtleties relating to the acquisition of the form.

We present the results of an experiment focusing on iconicity's role in the acquisition of forms. In line with Ortega & Morgan (2015) and Verhoef et. al. (2016), we predict that while iconicity helps to acquire new mappings, it may also lead to less precise encoding of forms. We presented learners (n = 36, no previous experience of a signed language) with an artificial gestural language based on iconic and arbitrary signs from BSL. We measured performance on an immediate imitation task, using the 3D body-tracking capabilities of Microsoft Kinect to quantify the trajectories of learners' wrists during production. This allows comparison of gestures produced by different participants using Dynamic

Time Warping (Celebi et. al. 2013). We also measured recall of the mapping using a guessing task, and recorded iconicity ratings on a 7 point Likert scale. Our results show that learners were indeed better at remembering mappings they perceived to be iconic (Figure 1), however, counter to our prediction, there was no difference in the accuracy with which they reproduced iconic and arbitrary items (Figure 2). A possible explanation for this is that the Kinect-based measured we used focuses on the trajectory of movement of the joints of the arms and wrists, whilst in the chosen stimuli iconicity was often based on hand-shape, rather than movement (e.g. in BSL 'TREE'). Future work will use stimuli that are more suited to the Kinect-based measure.



Acknowledgements

Asha Sato is funded by the Economic and Social Research Council (grant number ES/J500136/1).

gestures.

- Celebi, S., Aydin, A. S., Temiz, T. T., & Arici, T. (2013). Gesture Recognition using Skeleton Data with Weighted Dynamic Time Warping. In VISAPP (1) (pp. 620-625).
- Dingemanse, M., Blasi, D.E., Lupyan, G., Christiansen, M.H. and Monaghan, P., (2015). Arbitrariness, iconicity, and systematicity in language. *Trends in Cognitive Sciences*, 19(10), pp.603-615.
- Lockwood, G. and Dingemanse, M., (2015). Iconicity in the lab: A review of behavioral, developmental, and neuroimaging research into soundsymbolism. *Frontiers in psychology*, 6, p.1246.
- Perniss, P., Thompson, R. and Vigliocco, G., (2010). Iconicity as a general property of language: evidence from spoken and signed languages. *Frontiers in psychology*, 1, p.227.
- Ortega, G. and Morgan, G. (2015). Phonological development in hearing learners of a sign language: The influence of phonological parameters, sign complexity, and iconicity. *Language Learning*, 65(3), pp. 660-688.
- Verhoef, T., Kirby, S., & Boer, B. (2016). Iconicity and the emergence of combinatorial structure in language. *Cognitive science*, 40(8), 1969-1994.

SUCCESS IN SIGNALING: THE EFFECT OF FEEDBACK TO SIGNALER AND RECEIVER

JUSTIN SULIK $^{\ast 1}$ and GARY LUPYAN 2

*Corresponding Author: justin.sulik@gmail.com
¹Department of Psychology, Royal Holloway, University of London
²Department of Psychology, University of Wisconsin–Madison

Communicating without the benefit of a conventional signaling system is hard. Generating and interpreting novel signals involves complex inferences (Levinson, 1995), which often fail. One source of failure is a tendency not to take the other person's point of view, i.e., being egocentric (Sulik & Lupyan, under review). One solution to such difficulties is feedback (Barr, 2004; Spike, Stadler, Kirby, & Smith, 2017; Micklos, 2016), for example in the form of information about how the recipient behaved on receiving the signal. Here, we investigate how performance in a novel signaling task is affected by various kinds of feedback. We manipulated whether feedback is provided to signaler, receiver or both, and how informative the feedback is. The topic is relevant to language evolution because it helps us understand the factors that make complex communication possible when people cannot rely on the use of conventional word-meaning associations.

Pairs of participants took part in a signaling task (Sulik & Lupyan, 2016). Items, signals and guesses were single words. For instance, to get someone to guess 'bride', the signaler might generate the signal 'groom', which is quite likely to be successful. Alternatively, they might generate signal 'wedding' which is less likely to lead to a correct guess. Participants could generate any English word as a signal or guess (similar to the game show *Password*¹). Although the signals were English words, the overall task could not make full use of a shared conventional language because it required, for example, using 'groom' or 'wedding' to mean 'bride'. The target items were selected based on published word association norms (Nelson, McEvoy, & Schreiber, 2004), such that the participants' perspectives were aligned for some items ('symmetric' items) but not for others ('asymmetric' items).

The motivation for this unusual task is that it provides a clean measure of how informative signals are, all else being equal. In a more common gesture-based task, someone signaling 'moose' might gesture antlers, but performance is con-

¹http://bit.ly/2h1ZYK0

founded with the likelihood of recognizing that the gesture represents antlers, with the ease of gesturing those antlers, and with the degree of resemblance between the gesture and real antlers, which will vary between people.

Participants completed the task in pairs with each pair assigned to one of four conditions.

(1) *Full feedback*: the signaler was informed of the guesser's guess and the guesser was informed of the target item.

(2) *Minimal feedback*: both signaler and receiver received minimal feedback comprising only whether the guess was correct

(3) *Signaler-only feedback*: The signaler received full feedback; the receiver received no feedback.

(4) *Receiver-only feedback*: The receiver received full feedback; the signaler received no feedback.

Unsurprisingly, performance was better at symmetric items overall, though average performance was similar in each feedback condition. We next measured whether performance improved over the course of the game. To be clear: because participants saw each item once, we are not testing whether they discovered stable ways to signal those items through practice. Rather, our measure of improvement relates to discovering ways of succeeding at the task itself.

Overall, performance improved over time, but the improvement interacted with feedback and symmetry. For symmetric items, the improvement was independent of feedback. For asymmetric items, performance only improved when the signaler received full feedback (conditions 1 and 3). Performance in the minimal feedback and receiver-only feedback conditions (2 and 4) remained flat.

In sum, feedback had little effect when people's perspectives overlapped (symmetric items): performance improved regardless. However, when their perspectives differed, improvement at the task was driven by feedback to the signaler, though it was not enough for the signaler to just know whether they were correct or not. This investigation of factors that drive communicative success in the absence of a conventional signaling system helps us understand the prerequisites for the emergence of such a system.

- Barr, D. J. (2004). Establishing conventional communication systems: Is common knowledge necessary? *Cognitive Science*, 28(6), 937–962.
- Levinson, S. C. (1995). Interactional biases in human thinking. In Social intelligence and interaction (pp. 221–260). Cambridge University Press.
- Micklos, A. (2016). Interaction for facilitating conventionalization: Negotiating the silent gesture communication of noun-verb pairs. In S. Roberts, C. Cuskley, L. McCrohon, L. Barceló-Coblijn, O. Feher, & T. Verhoef (Eds.), *The evolution of language: Proceedings of the 11th international conference (EVOLANG11)*.

- Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida word association, rhyme, and word fragment norms. *Behavior Research Methods*, 36(3), 402–407.
- Spike, M., Stadler, K., Kirby, S., & Smith, K. (2017). Minimal requirements for the emergence of learned signaling. *Cognitive science*, *41*(3), 623–658.
- Sulik, J., & Lupyan, G. (2016). Failures of perspective taking in an open-ended signaling task. In S. Roberts, C. Cuskley, L. McCrohon, L. Barceló-Coblijn, O. Feher, & T. Verhoef (Eds.), *The evolution of language: Proceedings of the 11th international conference (EVOLANG11).*
- Sulik, J., & Lupyan, G. (under review). Perspective taking in a novel signaling task: effects of world knowledge and contextual constraint.

MODELING CREATIVITY AND COMMUNICATION

JUSTIN SULIK1

^{*}justin.sulik@gmail.com ¹Department of Psychology, Royal Holloway, University of London

Usually, humans use shared language to communicate successfully and efficiently. But people can also communicate — with considerably more effort and less predictable success — by creatively inventing novel signals. Modern humans need to do this comparatively rarely: when playing guessing games, or when interacting with people with whom they do not share a language. However, our prelinguistic ancestors had no shared words, so the creative generation of novel signals was one plausible route to conventional language.

This raises a cognitive puzzle. On one hand, the inferential mechanisms that underlie the creative generation of novel signals are computationally complex (Levinson, 1995, 2006; Stolk et al., 2013), and while any neurotypical human can learn language given the right input, performance at creative signaling is a great deal less uniform (Sulik & Lupyan, in prep.). On the other hand, the creative generation of novel signals was evolutionarily prior to the emergence of language.

Thus, creativity is a cognitively expensive bump in the road to language. The puzzle is how we got over that bump. Generating or interpreting a novel signal requires extensive investment of cognitive resources, and this investment does not predictably lead to communicative success. In this work I ask, 'in what situations do the computationally expensive mechanisms of creativity boost communicative success?' This question is tackled here with agent-based simulations of creativity and communication. This allows explicit manipulation of relevant features, such as agents' mental representations, communicative strategies, and context.

First, I review psychological accounts of the cognitive mechanisms that underlie creativity. For instance, Mednick (1962) characterizes creativity as involving a relatively flat associative hierarchy in semantic knowledge. Jung-Beeman (2005) argues that creative problem solving involves coarse coding (weak, broad spread of activation) as opposed to fine coding (strong, focused spread of activation). Other accounts emphasize the small-world network properties of creativity (Schilling, 2005; Kenett, Anaki, & Faust, 2014) or differences in executive control (Benedek et al., 2017). Some of these mechanisms are demonstrably related to novelty in communication (Mashal, Faust, Hendler, & Jung-Beeman, 2007).

Second, I introduce the model. Each agent is a graph with weighted, directed

edges (modeling a semantic network, with concepts as nodes and their associative relationships as edges). Agents interact repeatedly. In each interaction, one concept is chosen as the target to be communicated. The signaler generates a signal by sampling (probabilistically, based on edge weights) from the associates of the target in their semantic network, much like someone wanting to communicate 'moose' might gesture a semantic associate, antlers. The receiver makes a guess based on spreading activation in their semantic network when the signal nodes fire. The model explores the factors driving communicative success. There is no generational turnover or learning, though these are directions for future research.

Most of the model parameters describe the organization and structure of the agents' networks ('neural' parameters). Some of these are directly related to the above accounts of creativity (e.g. parameters governing the distribution of weights among a node's edges or the extent to which activation spreads), while others are not directly related to creativity, but do impact communicative success (e.g., how many neighbors a node has, or how similar the signaler and receiver's graphs are), and might thus be expected to interact with the creativity parameters. Other parameters ('context' parameters) describe the communicative context, such as how constrained the context is, or how complex the signal can be.

The parameters interact in complex ways to affect communication. The following are just some illustrations. It turns out that the complexity of the signal drastically alters how neural parameters drive success, and that signalers and receivers are often subject to different pressures. A flat associative hierarchy (cf. Mednick, above) makes signal generation more entropic, while a steep hierarchy decreases signaling entropy, but whether either of these boosts communicative success depends on a variety of other factors. For instance, creative, high-entropic signalers become less successful as the number of edges increases, but this effect is counteracted when the signaler's and receiver's representations become less similar, or when signals are allowed to become more complex. I introduce an R Shiny app to facilitate online interactive visual exploration of the parameter space.

While it may seem discouraging that no simple conclusion can be drawn about when creativity (or its opposite, rigidity) is adaptive, this result is actually informative about the evolution of language. Since the parameters interact in complex ways, then *if* the communicative context varies in complexity, it is adaptive to be able to flexibly employ *both* creative and rigid strategies. One solution is for individuals to have access to both creative and rigid computational styles; another way is to distribute the cognitive burden between individuals, so some are more creative and others more rigid. I review evidence suggesting that human social cognition makes use of both of these strategies (Jung-Beeman, 2005; Faust & Kenett, 2014; Sulik & Lupyan, in prep.). I propose that the first steps towards human language must have been characterized by such variation in communicative context, and an explanation of language evolution must thus encapsulate both social and cognitive divisions of labor.

- Benedek, M., Kenett, Y. N., Umdasch, K., Anaki, D., Faust, M., & Neubauer, A. C. (2017). How semantic memory structure and intelligence contribute to creative thought: a network science approach. *Thinking & Reasoning*, 23(2), 158–183.
- Faust, M., & Kenett, Y. N. (2014). Rigidity, chaos and integration: hemispheric interaction and individual differences in metaphor comprehension. *Frontiers in Human Neuroscience*, 8, 511.
- Jung-Beeman, M. (2005). Bilateral brain processes for comprehending natural language. *Trends in Cognitive Sciences*, *9*, 512-518.
- Kenett, Y. N., Anaki, D., & Faust, M. (2014). Investigating the structure of semantic networks in low and high creative individuals. *Frontiers in Human Neuroscience*, *8*, 407.
- Levinson, S. C. (1995). Interactional biases in human thinking. In Social intelligence and interaction (pp. 221–260). Cambridge University Press.
- Levinson, S. C. (2006). On the human "interaction engine". In N. J. Enfield & S. C. Levinson (Eds.), *Roots of human sociality: Culture, cognition and interaction* (pp. 39–69). Oxford: Berg.
- Mashal, N., Faust, M., Hendler, T., & Jung-Beeman, M. (2007). An fMRI investigation of the neural correlates underlying the processing of novel metaphoric expressions. *Brain and Language*, 100, 115-126.
- Mednick, S. A. (1962). The associative basis of the creative process. *Psychological Review*, 69(3), 220-232.
- Schilling, M. A. (2005). A 'small-world' network model of cognitive insight. *Creativity Research Journal*, 17(2-3), 121-154.
- Stolk, A., Verhagen, L., Schoffelen, J.-M., Oostenveld, R., Blokpoel, M., Hagoort, P., van Rooij, I., & Toni, I. (2013). Neural mechanisms of communicative innovation. *PNAS*, 110(36), 14574-14579.
- Sulik, J., & Lupyan, G. (in prep.). Individual differences in performance at a novel signaling task.

DEVELOPMENT OF VOCAL TEMPORAL PARAMETERS IN DISTANTLY RELATED VOCAL LEARNERS, BIRDS AND HUMANS

MIKI TAKAHASI*1,2, KAZUO OKANOYA3, and REIKO MAZUKA1,4

*Corresponding Author: mtakahas@brain.riken.jp
¹Lab. for Language Development, RIKEN BSI, Japan
²JSPS Research Fellow, Japan
³Department of Life Sciences, Graduate School of Arts and Sciences, the University of Tokyo, Japan
⁴Department of Psychology and Neuroscience, Duke University, Durham NC, USA

Young animals which are immature and immobile produce vocalizations such as isolation calls or cries to elicit care from adults. A common feature of these vocalizations is that the several vocal elements make a sequential vocalization. Sequential vocalizations accompany constant control of the air flow (reviewed in Doupe & Kuhl, 1999; Riede & Goller, 2010). Although sequential vocalizations eliciting care may largely be innate and involuntarily controlled for many animals, some animals, such as songbirds, are capable of learning new sequential vocalizations through social experiences. These vocal learners have been found to have a direct projection from cerebral cortex to vocal-respiratory center in medulla (Jurgens, 2002). This direct projection enables voluntary respiration and vocalization during vocal learning. Veit et al. (2011) showed juvenile songbirds learn the link between respiration and vocalization. In human infants, respiratory rate in crying and movements of lung and abdomen change through development (Boliek, Hixon, Watson, & Morgan, 1996; Wilder & Baken, 1974). However, it is not known yet when the shift from involuntary to voluntary respiration occurs in vocal learning process. The present paper analyzed the early vocalizations in human infants and Bengalese finches, one of songbird species, to address this question.

We focused on the developmental changes in 3 temporal parameters in sequential vocalizations: Note Duration (ND), Inter-Onset-Interval (IOI), and

Inter-Note-Interval, silent interval between two adjacent notes (INI). To produce sequential vocalizations, respiration must be controlled. Thus, the analyses of temporal parameters for sequential vocalizations may reveal the timing of the developmental shift from the involuntary to voluntary respiration control. We used the longitudinally recorded vocalization from 4 infants and 6 young finches. In infants data, crying at 0 month were also analyzed to compare with the sequential vocalizations.

For human infants, the distribution patterns of temporal parameters in cries differ from those in spontaneous sequential vocalizations as shown in Fig 1. The mean duration of ND gradually increased up to 5 months of age, when it became close INI. This pattern was similar to neonates' cries. In Bengalese finches, INI gradually declined up to 70 days, although merging of ND and INI occurred, too. However, combining a specific ND with a specific INI appeared through development in Bengalese finches. IOI kept its distribution pattern in humans. In general, the vocalizations of 5-month-olds are known to become highly variable as vocal playing. Based on these data, involuntary/voluntary switching might occur at 5 months of age in humans. In Bengalese finches, voluntary respiration control may enable to stabilize sequential vocalization at around 70 days.



Figure 1. Developmental change of mean durations of three temporal parameters.

Acknowledgements

This work was supported in part by JSPS Grant-in-Aid for JSPS Research Fellow (16J40180) to MT, Scientific Research S (16H06319) to RM, and MEXT Grant-in-Aid for Scientific Research on Innovative Areas in #4903 (Cocreative Language Evolution, 17H06380) to RM.

- Boliek, C. A., Hixon, T. J., Watson, P. J., & Morgan, W. J. (1996). Vocalization and breathing during the first year of life. *Journal of Voice*, 10(1), 1–22.
- Doupe, A. J., & Kuhl, P. K. (1999). BIRDSONG AND HUMAN SPEECH: Common Themes and Mechanisms. *Annual Review of Neuroscience*, 22(1), 567–631.
- Jurgens, U. (2002). Neural pathways underlying vocal control, 26.
- Riede, T., & Goller, F. (2010). Peripheral mechanisms for vocal production in birds - differences and similarities to human speech and singing. *Brain* and Language, 115(1), 69–80.
- Veit, L., Aronov, D., & Fee, M. S. (2011). Learning to breathe and sing: development of respiratory-vocal coordination in young songbirds. *Journal of Neurophysiology*, 106(4), 1747–1765.
- Wilder, C. N., & Baken, R. J. (1974). Respiratory patterns in infant cry. *Human Communication*, (Winter), 18–34.

DO SONGBIRDS PERCEIVE SYLLABLES AS ISOLATED TOKENS?

TOMOKO MIZUHARA¹ and KAZUO OKANOYA^{*1}

*Corresponding Author: cokanoya@mail.ecc.u-tokyo.ac.jp ¹Department of Life Sciences, The University of Tokyo, Tokyo, Japan

The songbird is an established model for comparative biology of human speech acquisition (Marler, 1970; Doupe & Kuhl, 1999). Many attempts to explore further parallels have been made to date: some described song syntax of Bengalese finches (e.g. Okanoya, 2004) and others trained zebra finches in artificial grammar composed of syllables (e.g. van Heijningen *et al.*, 2013). Here they share a tacit assumption — songbirds consider syllables as tokens.

It is a common practice for researchers in this field to isolate each syllable from birdsong for analysis. However, songbirds basically do not vocalize individual syllables apart from song bouts. In addition, Dooling & Searcy (1980) pointed out that auditory masking effects may not be ignorable when song notes are perceived.

We hypothesized that if songbirds perceived syllables as tokens, they would detect a syllable regardless of temporal distance from another syllable. 6 female Bengalese finches were first trained to detect syllable [a] that were coupled with another syllable [p]. In p-t condition [p] preceded [a]; in t-p condition [p] followed [a]. The interval had 4 variants that were uncommon in Bengalese finch song (Figure 1A). After reaching a predetermined pass criterion, probe stimuli were introduced. They had shorter or longer interval than that of training stimuli (Figure 1A).

Birds with odd ID went through p-t condition first, and birds with even ID went the other way around (Figure 1B). Difference in mean days each group took for the first condition was not significant (Welch's two sample t-test, t = 4.05, df = 2, P = 0.055). Once they finished either condition, they did not require so long days to get accustomed to new conditions (paired t-test, t = 5.87, df = 5, P = 0.0020).

Figure 1C shows that response rate decreased as interval was shortened in p-t condition except bird#4, but not in t-p condition. We fitted data of each condition with generalized linear mixed models (GLMMs). The null model (the effect of log(gap length) equals 0) was rejected in p-t condition (P< 0.001) but not in t-p condition (P = 0.47) with likelihood ratio test. This result is consistent with the fact that effect of forward masking is greater than that of backward masking (Dooling & Searcy, 1980). In other words, tokenization was inhibited in p-t condition in spite of the training with variable intervals.

Overall, it is not taken for granted that songbirds regard syllables as tokens unconditionally. We do not think that the brief interval itself was distracting for birds: If so, their response rate would have shown some decline in t-p condition, too. Response rate in t-p condition, however, kept high independent of the stimulus type. Moreover, in p-t condition, the longer the interval was, the higher the response rate was. These raises the possibility that tokenization might be helped by elongating inter-syllable interval. Such speculation is also consistent with the result of training days comparison between before and after switching, as birds, especially with odd ID (p-t first), soon passed their novel condition. We hope that our study will benefit future artificial grammar learning studies not only in songbirds, but also in other animals including humans.



Figure 1 (A) Distribution of gap length in Bengalese finch song (data from 8 males). Numbers below represent length of silent intervals adopted in this study. A black bar indicates 4 variants for training stimuli. (B) Training days required for discrimination. (C) Response rate to probe stimuli.

Acknowledgements

We thank Dr. Gabriël J. L. Beckers for his comments on the experimental plan, Dr. Takuya Koumura for gap distribution data of Bengalese finch song, Ms. Tomoko Fujii for providing song data to make stimuli, and Ms. Yoshiko Sawada for her cooperation in bird care. This study was supported by MEXT Innovative Areas #4903, JP17H06380 to K.O. and by the University of Tokyo Grant for PhD Research RA to T.M..

- Dooling, R. J., & Searcy, M. H. (1980). Forward and backward auditory masking in the parakeet (Melopsittacus undulatus). *Hearing research*, 3(4), 279-284.
- Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Annual review of neuroscience*, 22(1), 567-631.
- Marler, P. (1970). Birdsong and Speech Development: Could There Be Parallels? There may be basic rules governing vocal learning to which many species conform, including man. *American Scientist*, 58(6), 669-673.
- Okanoya, K. (2004). Song syntax in Bengalese finches: proximate and ultimate analyses. *Advances in the Study of Behavior*, *34*, 297-346.
- van Heijningen, C. A., Chen, J., van Laatum, I., van der Hulst, B., & ten Cate, C. (2013). Rule learning by zebra finches in an artificial grammar learning task: which rule? *Animal cognition*, 16(2), 165-175.

DIFFERENCE OF TWO RECURSIONS, "HIERARCHICAL EMBEDDING" AND "SELF-REFERENCE": FROM VIEW OF ADAPTIVE FUNCTIONS AND IMPLEMENTATIONS

GENTA TOYA*1, RIE ASANO2, and TAKASHI HASHIMOTO1

*Corresponding Author: toyagent@jaist.ac.jp ¹School of Knowledge Science, Japan Advanced Institute of Science and Technology, Nomi, Japan ²Department of Systematic Musicology, University of Cologne, Cologne, Germany

"Recursion" is understood in two ways: 1) "hierarchical embedding" as a structure and 2) "self-reference / self-invocation of a function" as an operation. The former refers to structure itself, and the latter refers to an operation¹ to create the structure. In research of language and comparative cognition between humans and non-human animals, these two kinds of "recursion" were sometimes confused and misunderstood (Martins, 2012). In the current talk, we demonstrate that the conceptual essence of recursive operation is not the existence and generation of a certain hierarchical embedded structure, but a function which calls itself to create a new embedded structure indefinitely. Additionally, we suggest that the ecological environment of humans is likely to have made the latter adaptive, and the human brain has neural infrastructure that realizes recursive operation. Finally, from research on bird songs, human uniqueness of the recursive operation is discussed.

Goal-directed behavior in human beings includes not only non-recursive operation which is generally observed in non-human animals but also recursive operation in which parts of sequence are maintained so that it can be combined with other parts of sequence and arranged to have a particular form. However, the recursive operation requires a higher cognitive load to hold temporal processes. An evolutionary simulation study showed that the adaptive significance of self-

¹ In our previous paper (Toya and Hashimoto, 2017), this operation is called "recursive combination operation."

reference in object manipulation is to produce complex and diverse products such as stone tools (Toya and Hashimoto, 2015) in a highly competitive environment for resources (Toya and Hashimoto, 2017). Such environment may be realized from 2.5 million years ago with forest lands decreasing in the human evolution (Bobe and Behrensmeyer, 2004; deMenocal, 2011). The results of those simulations showed that recursive operation is required to achieve diverse goals although non-recursive operation has a lower cost than recursive operation. This consideration opens up a new perspective to relate recursive operation to a neural mechanism.

We hypothesize that recursive operation is implemented in the brain by focusing on the functionary segregated cortico-basal ganglia-thalamocortical (CBGT) circuits. The CBGT circuits are suggested to be necessary for working memory (Lustig, 2005; Hochstadt, 2006) and cognitive control (Monchi et al., 2001; Lieberman, 2002) in which maintenance and manipulation of temporal events play a significant role. Both two functions are crucial for realizing recursive operation.

Does a neural mechanism for recursive operation exist in other animals? In bird songs, basic elements are combined into chunks, which combined into phrases, which combined into whole sequence. Songbirds learn and use processing rules of these songs in separate brain modules for each level of the hierarchy (Okanoya, 2004; Okumura et al., 2011). This mechanism seems to be different from the recursive operation because songbirds do not generate new structure which do not belong to learned set. Therefore, we suggest that recursive operation is not required for reproducing hierarchical embedding of learned songs.

In conclusion, humans possess the ability of "recursive operation" with selfreference, which leads to creativity that is not found in the hierarchical embedding for learning-reproduction system such as the song of songbirds.

References

- Berwick, R. C., Okanoya, K., Beckers, G. J. L., & Bolhuis, J. J. (2011). Songs to syntax: the linguistics of birdsong. *Trends in Cognitive Sciences*, 15(3), 113– 121.
- Bobe, R., & Behrensmeyer, K. (2004). The expansion of grassland ecosystems in Africa in relation to mammalian evolution and the origin of the genus *Homo*. *Palaeogeography Palaeoclimatology Palaeoecology*, 207, 399-420.

deMenocal, P. B. (2011). Climate and human evolution. Science, 331, 540-542.

Hochstadt, J., Nakano, H., Lieberman, P., & Friedman, J. (2006). The roles of sequencing and verbal working memory in sentence comprehension deficits in Parkinson's disease. *Brain and Language*, 97(3), 243–257.

- Lieberman, P. (2002). On the nature and evolution of the neural bases of human language. *American Journal of Physical Anthropology*, *119*(*S35*), 36–62.
- Lustig, C., Matell, M. S., & Meck, W. H. (2005). Not "just" a coincidence: frontalstriatal interactions in working memory and interval timing. *Memory*, 13(3– 4), 441–448.
- Monchi, O., Petrides, M., Petre, V., Worsley, K., & Dagher, A. (2001). Wisconsin Card Sorting revisited: distinct neural circuits participating in different stages of the task identified by event-related functional magnetic resonance imaging. *Journal of Neuroscience*, 21(19), 7733–41.
- Martins, M. D. (2012). Distinctive signatures of recursion. *Philosophical Transactions of the Royal Society B*, 367(1598), 2055–2064.
- Okanoya, K. (2004). The Bengalese finch: A window on the behavioral neurobiology of birdsong syntax. Annuals of the New York Academy of Sciences, 1016, 724–735.
- Okumura, T., Yamashita, Y., Okanoya, K., & Tani, J. (2011). Songbirds as experimental model for studying complex sequential behaviors and sensorymotor learning. *Journal of Japanese Neural Network Society*, 18(3), 135– 146.
- Toya, G., & Hashimoto, T. (2015). Computational study on evolution and adaptability of recursive operations. *Proceedings of the 20th International Symposium on Artificial Life and Robotics* (pp. 68–73). (CD-ROM).
- Toya, G., & Hashimoto, T. (2017). Evolution of recursive combination operation. In C. Knibbe, G. Beslon, D. Parsons, D. Misevic, J. Rouzaud-Cornabas, N. Bredéche, S. Hassas, O. Si-monin, & H. Soula (Eds.), *Proceedings of The European Conference on Artificial Life 2017* (p. 396–403). Cambridge, MA: The MIT Press.

COMMUNICATIVE ASYMMETRY, HANDEDNESS AND LANGUAGE EXPERIENCE

OLGA VASILEVA

ovasilev@sfu.ca

Department of Psychology, Simon Fraser University, Vancouver, Canada

Lateralization of human cognitive and linguistic functioning has important implications for the evolution of our species. Various researchers suggest that a link between communicative and manual development in humans, in both ontogeny and phylogeny, led to a human-specific brain lateralization pattern for language, tool-use and possibly other cognitive domains (Corballis, 2003; Fitch & Braccini, 2013). The human-unique character of these asymmetries (namely, the fact that humans are the only animals demonstrating such strong population-level biases in both asymmetries), and their presumed importance for human-specific traits, such as tool-use and language (e.g. Estalrrich & Rosas, 2013; Chance & Crow, 2007), have long attracted researchers' attention and became a field of studies in its own right.

However, the exact nature of handedness-language relationship, as well as its evolutionary history is far from being understood. Particularly concerning is the fact that studies conducted on human infants and primates bring inconsistent results. Some researchers (e.g. Fagard, 2013) have expressed serious concerns regarding the level of understanding of manual/communicative asymmetries relations in human ontogeny and phylogeny. Although there is evidence suggesting that language and handedness are associated, more research is required to test this hypothesis.

There are two important notions regrading such a research. First of all, it is necessary to investigate the association between language asymmetry, handedness and other cognitive parameters (Cochet, 2016). Second of all, it is necessary to conduct studies on more diverse samples. The majority of studies investigating manual and communicative asymmetries have been conducted with either human infants or non-human primates, while the number of studies addressing these aspects in the adults is quite limited. Such situation is problematic for our ability to relate the findings of research in children and primates with the research conducted on adult population. Moreover, the studies

addressing the asymmetries in the adult population, as a rule, rely on samples with a high level of homogeneity (e.g. age, language background or handedness). However, an integrated approach to the problem of manuallinguistic asymmetries relationship demands not only comparative, but a more comprehensive developmental data, and that is, examination of manual handedness and language not only in children, but necessarily in the adults in more diverse samples. To this day, only a limited number of studies have employed such approach.

The proposed submission outlines a project attempting to apply such an integrative approach and fill in existing gaps in the research. The purpose of the project is to investigate the relationship between manual and linguistic asymmetries relation participants' in to the language experience (bilinguals/monolinguals). According to two recent meta-analyses bilinguals exhibit a more balanced asymmetry in language processing with the right hemisphere being significantly involved in language processing (Vaid & Hull, 2006; Vaid & Hull, 2007). Moreover, the results indicate that overall monolinguals (MG) and late bilinguals (LB) (individuals acquiring a second language after the age of six) demonstrate a more left-lateralized pattern of language processing, while the early bilinguals (EB) (individuals acquiring a second language before the age of six) are less lateralized in that respect. Consequently, since BL differ from ML in language asymmetry, and since there is a presumed connection between manual and language asymmetry in the adults (Cochet & Vauclair, 2012), it might be expected that bilinguals also differ from monolinguals in handedness.

No previous research has addressed the question of handedness – language asymmetry in relation to bilingualism. Additionally, while many previous studies have employed indirect measures of language asymmetry (e.g. asymmetry of a communicative gesture), in the present study participants will take a dichotic listening test to determine their actual linguistic asymmetry profile (Hugdal, 2003). If language asymmetry and handedness is associated in the adults, we can expect congruent relationship between these parameters in each of the groups (EB, LB, ML) and significant differences between groups.

Finally, in order to capture the relationship between handedness and language more effectively, the study would employ additional measures allowing separating these relationships from potential co-founding parameters, such as general motor asymmetry (footedness) and cognitive factors (autism-like and schizotypy traits measured in non-clinical populations and previously associated with mixed handedness and decreased lateralization (e.g. Sommer et al., 2009).

The studies are using established methodology on novel populations which allows comparing results with the existing findings in the field and, at the same time, broadening a research area of asymmetry development and its implications for the evolution of cognition and language in humans.

- Chance, S. A., & Crow, T. J. (2007). Distinctively human: cerebral lateralisation and language in Homo sapiens. J Anthropol Sci, 85, 83-100.
- Cochet, H. (2016). Manual asymmetries and hemispheric specialization: Insight from developmental studies. Neuropsychologia, 93, 335-341.
- Cochet, H., & Vauclair, J. (2012). Hand preferences in human adults: Noncommunicative actions versus communicative gestures. cortex, 48(8), 1017-1026.
- Corballis, M. C. (2003). From hand to mouth: The origins of language. Princeton University Press.
- Estalrrich, A., & Rosas, A. (2013). Handedness in Neandertals from the El Sidrón (Asturias, Spain): evidence from instrumental striations with ontogenetic inferences. PLoS One, 8(5), e62797.
- Fagard, J. (2013). Early development of hand preference and language lateralization: are they linked, and if so, how?. Developmental psychobiology, 55(6), 596-607.
- Fitch, W., & Braccini, S. N. (2013). Primate laterality and the biology and evolution of human handedness: a review and synthesis. Annals of the New York Academy of Sciences, 1288(1), 70-85.
- Hugdahl, K. (2003). Dichotic listening: an experimental tool in clinical neuropsychology. In Experimental methods in neuropsychology (pp. 29-46). Springer US.
- Hull, R., & Vaid, J. (2007). Bilingual language lateralization: A meta-analytic tale of two hemispheres. Neuropsychologia, 45(9), 1987-2008.
- Somers, M., Sommer, I. E., Boks, M. P., & Kahn, R. S. (2009). Hand-preference and population schizotypy: a meta-analysis. Schizophrenia research, 108(1), 25-32.
- Vaid, J., & Hull, R. (2006). Laterality and language experience. Laterality, 11(5), 436-464

COGNITIVE BIASES AND CULTURAL EVOLUTION IN THE EMERGENCE OF SPACE-TIME MAPPINGS IN LANGUAGE

TESSA VERHOEF^{*1,2}, ESTHER WALKER³, and TYLER MARGHETIS⁴

*Corresponding Author: tverhoef@ucsd.edu

¹Leiden Institute of Advanced Computer Science, Leiden University, Leiden, The Netherlands

²Center for Research in Language, University of California, San Diego, La Jolla, USA ³Cognitive Science, University of California, San Diego, La Jolla, USA ⁴Psychological and Brain Sciences, Indiana University, Bloomington, USA

Humans spatialize time. This occurs in artifacts like timelines, in gesture, and — critically — in conventional language ("look FORWARD to the future"). This phenomenon, however, exhibits considerable cross-linguistic variability. Explaining the origins of space-time mappings in language, therefore, will require considering both cultural processes and individual biases (Núñez & Cooperrider, 2013). Here we present two laboratory experiments on the interplay of these mechanisms in the emergence of space-time mappings.

In our first experiment, English-speaking pairs of participants (i.e., dyads) played a guessing game (cf. Garrod et al., 2007; Healey et al., 2007; etc.) about temporal concepts, using a novel, spatial signaling device: a vertical bar with a moving bubble controlled through a touch screen (Fig. 1a). After repeated interaction and social coordination, dyads developed shared signals that used space to express aspects



Figure 1: (a) Signaling device, (b) Consistent use of spatial length to communicate relative duration, (c) and (d) Contrasting mappings for past/future.

of time. Some of these conventions were nearly universal. For instance, every dyad used greater spatial *lengths* to indicate greater temporal *durations* (Fig. 1b). Conventions using vertical location to differentiate past from future also recurred across dyads, but they differed in their precise mapping (e.g., up = past or up = future, (Fig. 1c, d)). This mix of *shared* and *idiosyncratic* conventions suggests the involvement of both shared biases and cultural processes. While all dyads developed space-time mappings, which facilitated communication, these were never sufficiently elaborated or systematic to achieve perfect communication.

Our second experiment tested how *transmission* would further shape the signalling systems that emerged in the guessing game from Experiment 1. We hypothesized that iterated transmission across generations of interacting users might prompt the evolution of more regular and stable systems (Tamariz et al, 2012). The signals that were



developed by one dyad were used as initial training for the next dyad, before they started interacting. We tested six chains of eight dyads. After repeated transmission, some dyads achieved near-perfect communicative success, as a result of the gradual appearance of more fully systematized space-time mappings. Figure 2 illustrates an emerged system, which evolved highly systematic mappings and compositionality.

These laboratory experiments illustrate how space-time mappings in language can evolve from interacting mechanisms involving individual biases, social coordination and iterated transmission. The interplay between these mechanisms may explain the cross-linguistic *commonalities* as well as the *variety* found in space-time mappings.

- Garrod, S., Fay, N., Lee, J., Oberlander, J. & MacLeod, T. (2007). Foundations of representation: Where might graphical symbol systems come from? *Cognitive Science*, 31(6): 961–987.
- Healey, P. G. T., Swoboda N., Umata I. & King J. (2007). Graphical language games: interactional constraints on representational form. *Cognitive Science*, 31: 285–309.

- Núñez, R. & Cooperrider, K., (2013). The Tangle of Space and Time in Human Cognition. *Trends in Cognitive Sciences*, 17(5): 220-229.
- Tamariz, M., Cornish, H., Roberts, S. & Kirby, S. (2012) The effect of generation turnover and interlocutor negotiation on linguistic structure. In *Proceedings of EVOLANG9*. World Scientific, (pp. 555-556).

VOCAL LEARNING IN BATS: FROM GENES TO BEHAVIOUR

SONJA C. VERNES *1,2

 *Corresponding Author: sonja.vernes@mpi.nl
 ¹ Neurogenetics of Vocal Communication Group, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
 ² Donders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands

1. Vocal learning

Vocal production learning (herein 'vocal learning') - the ability to modify vocal signals based on conspecific auditory feedback - is an essential component of human spoken language (Bolhuis, Okanoya, & Scharff, 2010a; Nowicki & Searcy, 2014). Infants learning to speak must perform this task, which includes auditory perception, memorization of template, vocal motor planning and production, template matching to determine goodness of fit and modification of the vocal-motor output as necessary (Petkov & Jarvis, 2012). The complexity of this task suggests that multiple neurobiological and genetic mechanisms are likely to underlie its evolution and biological encoding. Given the necessary limitations of studying vocal learning in humans, animal models represent an opportunity to understand the neurogenetic mechanisms underlying this spoken language-relevant trait.

1.1. Animal models of vocal learning

Vocal learning is a rare trait in the animal kingdom. Few non-human species have convincingly shown this trait. To date, vocal learning has been documented in some cetacean, pinniped, elephant, bat and bird species (Bolhuis, Okanoya, & Scharff, 2010b; Janik & Slater, 1997). Songbirds have dominated the study of vocal learning due to their well-defined learning paradigm, sexual dimorphism of their song, their ease of handling and ability to breed in captivity (Bolhuis et al., 2010b; Condro & White, 2014). Mammalian vocal learning has, by comparison, been understudied. Sea mammal vocal learning has been most well documented

from this group (Janik & Slater, 2000; Janik & Slater, 1997), however due to their size and habitat they present hurdles for neurological or genetic interrogations.

2. Bat models of vocal learning

Bats are highly social animals that have developed sophisticated vocal systems for navigation and communication (Vernes, 2017). Social communication in bats is often facilitated by low frequency calls (in the hearing range) and in some species these calls show evidence that they may be learned (Knörnschild, 2014). Two species with evidence for learned calls, *Phyllostomus discolour* and *Rousettus aegyptiacus*, can be maintained in laboratory colonies, making them amenable to neurogenetic manipulations (Esser, 1994; Knörnschild, 2014; Prat, Taub, & Yovel, 2015). Thus, their small size, ability to breed in captivity and sophisticated vocal communications, make bats an exciting mammalian model for the study of vocal learning.

2.1. Studying vocal learning behaviour in bats

It is essential to have a well-defined and highly controlled behavioural paradigm to dissect out the neuro-genetic mechanisms underlying vocal learning. We are developing a vocal learning paradigm in which bats are induced to learn novel vocalisations via an automated playback design. To date volitional control of vocal production and vocal plasticity (usage learning) has been demonstrated.

2.2. Studying the neurogenetics of vocal learning in bats

Multiple approaches are being used to understand the neurogenetics contributing to vocal behaviour. Activation of neurons in response to perception or production of communication calls is being mapped to determine the brain regions involved. To facilitate evolutionary comparisons and molecular genetic studies, *de novo* sequencing of *P. discolour* and *R. aegyptiacus* genomes are being performed. To understand the role of language-related genes, the expression patterns of *FoxP2*, *FoxP1* and *CntnaP2* have been comprehensively mapped in the brains of both species. Finally, gene knockdown studies will determine the contribution of genes such as *FoxP2* to vocal behavior.

Bringing together genetic, neurobiological and behavioural studies in this way will shed light on the encoding of vocal learning in bats, and ultimately inform our understanding of the evolution of this language-relevant trait.

Acknowledgements

This work was funded by a Marie Curie Career Integration Grant (PCIG12-GA-2012-333978), and a Max Planck Research Group Grant both awarded to S.C.V. This work was also funded by a Human Frontiers Science Program (HFSP) Research grant (RGP0058/2016) and is part of a collaborative project involving Prof. Lutz Wiegrebe (LMU, Germany), Prof Michael Yartsev (UC Berkeley, USA) and Prof Uwe Firzlaff (TUM, Germany). With thanks to Ella Lattenkamp, Janine Mengede, Ine Alvarez van Tussenbroek, Laura Baas and Ksenia Lavrichenko for their contributions to this project.

- Bolhuis, J. J., Okanoya, K., & Scharff, C. (2010a). Twitter evolution: converging mechanisms in birdsong and human speech. *Nat Rev Neurosci*, 11(11), 747-759. doi:10.1038/nrn2931
- Condro, M. C., & White, S. A. (2014). Recent Advances in the Genetics of Vocal Learning. *Comparative cognition & behavior reviews*, 9, 75-98. doi:10.3819/ccbr.2014.90003
- Esser, K. H. (1994). Audio-vocal learning in a non-human mammal: the lesser spear-nosed bat Phyllostomus discolor. *Neuroreport*, 5(14), 1718-1720.
- Janik, V. M., & Slater, P. J. (2000). The different roles of social learning in vocal communication. *Animal behaviour*, 60(1), 1-11. doi:10.1006/anbe.2000.1410
- Janik, V. M., & Slater, P. J. B. (1997). Vocal learning in mammals. Advances in the Study of Behavior, Vol 26, 26, 59-99. doi:Doi 10.1016/S0065-3454(08)60377-0
- Knörnschild, M. (2014). Vocal production learning in bats. Current opinion in neurobiology, 28, 80-85. doi:10.1016/j.conb.2014.06.014
- Nowicki, S., & Searcy, W. A. (2014). The evolution of vocal learning. *Current* opinion in neurobiology, 28, 48-53. doi:10.1016/j.conb.2014.06.007
- Petkov, C. I., & Jarvis, E. D. (2012). Birds, primates, and spoken language origins: behavioral phenotypes and neurobiological substrates. *Frontiers in evolutionary neuroscience*, 4, 12. doi:10.3389/fnevo.2012.00012
- Prat, Y., Taub, M., & Yovel, Y. (2015). Vocal learning in a social mammal: Demonstrated by isolation and playback experiments in bats. *Science Advances*, 1(2), e1500019. doi:10.1126/sciadv.1500019
- Vernes, S. C. (2017). What bats have to say about speech and language. *Psychonomic Bulletin & Review*, 24(1), 111-117. doi:10.3758/s13423-016-1060-3

HONEYBEES LEARN PATTERNS WITH FLASHING LIGHTS

SARA PERUFFO¹, ELISABETTA VERSACE $^{\ast1,2},$ BETTINA MACCAGNANI², and GIORGIO VALLORTIGARA²

*CORRESPONDING AUTHOR: E.VERSACE@QMUL.AC.UK

¹Center for Mind/Brain Sciences, University of Trento, Rovereto, Italy ²Department of Biological and Experimental Psychology, Queen Mary University of London, UK

1. Introduction

From linguistic stimuli to spatial and causal relations, extracting regularities is necessary to make sense of stimuli available in the world. Abilities to process abstract regularities have been observed across different species of mammals (Endress, Carden, Versace, & Hauser, 2010; Murphy, Mondragón, & Murphy, 2008) and birds (Spierings & ten Cate, 2016; Versace, Regolin, & Vallortigara, 2006; Versace, Spierings, Caffini, ten Cate, & Vallortigara, 2017). An open question is whether insects exhibit pattern learning abilities. Honeybees are an ideal candidate to address this issue because they exchange with conspecifics complex information through the honeybee dance, and master abstract concepts such as "same" and "different" (Giurfa, Zhang, Jenett, Menzel, & Srinivasan, 2001). Little is known, though, on their capacities to extract patterns from serially presented stimuli. Here we investigate the capacities of the honeybee *Apis mellifera* to extract temporal patterns by discriminating between lights presented with different rhythms. We trained foragers to gain a sucrose reward by choosing a flashing *vs.* a static light and documented honeybees' capacity to solve this discrimination task.

2. Methods

Individually marked bees (*Apis mellifera*) were trained to run through a Y-maze to collect 50% (weight/weight) sucrose solution. The apparatus, covered with a transparent net, was located in front of a window and illuminated by daylight. Bees
entered the apparatus through the open window and found at the end of the corridor two blue circles (6 cm diameter) illuminated by LED lightbulbs, one on the right and one on the left branch (see Fig. 1). One stimulus was illuminated by a static light, the other by a flashing light. For each subject, only the flashing (static) stimulus was rewarded with sucrose solution, while the other stimulus was baited with a bitter, quinine solution. The position of the rewarded stimulus was randomized between trials.



Figure 1. Overhead view of the apparatus.

We analyzed the individual and overall performance of bees that made a choice in the apparatus at least 24 times, using a two-tailed binomial test for individual performance, and a two-tailed one-sample Mann-Whitney test against the chance level for the overall performance.

3. Results and discussion

Seven bees made at least 24 choices in the apparatus. Four out of seven bees reached a performance significantly different from chance (18 or more correct responses: 21/24, 87.5%, p<0.001; 20/24, 83% p<0.01; 18/24, 75%, p=0.023; 18/24, 75%, p=0.023), while three bees did not (17/24, 71%, p=0.064; 14/24, 58%, p=0.541; 10/24, 42%, p=0.541). All bees with a significant performance chose the rewarded stimulus. Overall, we observed a significant preference for the rewarded stimulus (V=27, p=0.034), in fact all bees except one choose the rewarded pattern more often than the other one. These results suggest that honeybees can discriminate between static and flashing lights and that these insects are a suitable model for further investigation of artificial grammar learning. To the best of our knowledge, this is the first example of temporal pattern learning documented in honeybees. Insects can shed light into the evolution of pattern learning that does not depend on language.

- Endress, A. D., Carden, S., Versace, E., & Hauser, M. D. (2010). The apes' edge: positional learning in chimpanzees and humans. *Animal Cognition*, 13(3), 483–95. https://doi.org/10.1007/s10071-009-0299-8
- Giurfa, M., Zhang, S., Jenett, A., Menzel, R., & Srinivasan, M. V. (2001). The concepts of "sameness" and "difference" in an insect. *Nature*, 410(6831), 930–933. https://doi.org/10.1038/35073582
- Murphy, R. A., Mondragón, E., & Murphy, V. A. (2008). Rule learning by rats. Science, 319(5871), 1849–51. https://doi.org/10.1126/science.1151564
- Spierings, M. J., & ten Cate, C. (2016). Budgerigars and zebra finches differ in how they generalize in an artificial grammar learning experiment. *Proceedings of the National Academy of Sciences*, 3977–3984. https://doi.org/10.1073/pnas.1600483113
- Versace, E., Regolin, L., & Vallortigara, G. (2006). Emergence of Grammar as Revealed by Visual Imprinting in Newly-hatched Chicks. In: The Evolution of Language. In *Proceedings of the 6th International Conference, Rome, 12-15 April 2006.*
- Versace, E., Spierings, M. J., Caffini, M., ten Cate, C., & Vallortigara, G. (2017). Spontaneous generalization of abstract multimodal patterns in young domestic chicks. *Animal Cognition*. https://doi.org/DOI 10.1007/s10071-017-1079-5

RAPID TURN-TAKING AS A CONSTANT FEATURE OF SIGNED CONVERSATIONS

CONNIE DE VOS 1*1

*Corresponding Author: c.devos@let.ru.nl ¹Center for Language Studies, Radboud University, Nijmegen, the Netherlands

Social interaction is the primary ecological niche for languages to evolve and to emerge (Levinson 2006). The spontaneous conversations that characterize our every-day interactions are remarkably rapid across a wide range of typologically diverse spoken languages with most turns timed ~200ms after the prior (Stivers et al. 2009). Corpus analysis of the Sign Language of the Netherlands (NGT) has shown that this main observation also holds in the visual-gestural language modality, when the different movement phases of the hands are taken into account (de Vos et al. 2015). The present study is the first to investigate whether turn-timing is a constant pressure of conversation in language emergence by looking at different generations of the emergent signing variety Kata Kolok.

Kata Kolok is a signed language that has emerged outside the influence of any other signed languages. It has been used by six subsequent generations of native signers in a village community of Bali. On a par with the Stivers et al. comparative study, I collected question-answer sequences from the Kata Kolok Corpus (de Vos 2016), which covers generations III-V of adult signers in dyadic and triadic settings, as well as the NGT Interactive Corpus (de Vos et al. 2015). Prior work on sign languages shows that signed turns at talk can be effectively delineated on the basis of the lexically-specified movement phases of signs as well as prosodic turn-boundary markers such as eye blinks and nods. Each question-answer sequence was therefore transcribed for their movement phases (preparation, stroke, hold, retraction) as well as prosodic cues marking questionhood and phrase boundaries (e.g. blinks, raised eyebrows, nods) (cf. Kita et al. 1999). Turn-timing was calculated stroke-to-stroke, i.e. as the floor transition offset (FTO) between the end of the question's turn-final stroke and the start of the answer's turn-initial stroke (cf. de Vos et al. 2015). I report on both overall means for stroke-to-stroke turn transitions as well as the "Optimal turntiming" measure, which subtracts the cross-linguistic default of 200ms off the original FTO and collapses them into an absolute value to assess how effective

turn transitions are across generations as well as both language communities. Linear mixed effects modelling in R reveals there are no significant effects of either language community or generation, but does indicate that NGT signers are faster in triadic settings when allowing random effects for recording session and signer. This is illustrated by Figure 1. below.



Figure 1. Floor transition offsets in Kata Kolok and Sign Language of the Netherlands (NGT)

The present results further extend the generalization that signed languages are within the cross-linguistic range that can be expected on the basis of spoken language turn-timing, by including data from the sign language isolate Kata Kolok (cf. Stivers et al. 2009). Moreover, these results hold across generations suggesting that rapid turn-timing might a stable aspect of conversation in language emergence. The fact that NGT signers take turns more efficiently in triadic settings is consistent with the view that rapid turn-taking exerts a time pressure on both language producers and perceivers in such settings.

Prior work has shown that prosodic cues (e.g. of questionhood) are crucial in the anticipation of turn-endings (Casillas et al. 2015). Moreover, such prosodic signaling appears to be unstable in the early stages of sign language emergence (Sandler et al. 2011). Future work therefore focuses on the hypothesis that the time-pressure of real-life turn-taking may lead to the optimization of the prosodic features of utterances during sign language emergence.

Acknowledgements

This work was supported by the NWO VENI grant 'The face in sign language interaction' awarded to Connie de Vos.

- Casillas, M., De Vos, C., Crasborn, O., & Levinson, S. C. (2015). The perception of stroke-to-stroke turn boundaries in signed conversation. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. R. Maglio (Eds.), Proceedings of the 37th Annual Meeting of the Cognitive Science Society (CogSci 2015) (pp. 315-320). Austin, TX: Cognitive Science Society.
- De Vos, C. (2016). Sampling shared sign languages. *Sign Language Studies*, 16(2), 204-226.
- De Vos, C., Torreira, F., & Levinson, S. C. (2015). Turn-timing in signed conversations: Coordinating stroke-to-stroke turn boundaries. Frontiers in Psychology, 6: 268.
- Kita, S., Van Gijn, I., and van der Hulst, H. (1998). "Movement phases in signs and co-speech gestures, and their transcription by human coders," in *Gesture and Sign Language in Human-Computer Interaction*, eds I. Wachsmuth and M. Fröhlich (Berlin: Springer), 23–35.
- Levinson, S. C. (2006). "On the human interaction engine," in *Roots of Human Sociality: Culture, Cognition and Interaction*, eds N. J. Enfield and S. C. Levinson (Oxford: Berg), 39–69.
- Sandler, W., Meir, I., Dachkovsky, S., Padden, C., & Aronoff, M. (2011). The emergence of complexity in prosody and syntax. *Lingua*, 121(13), 2014-2033.
- Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., et al. (2009). Universals and cultural variation in turn-taking in conversation. *Proc. Natl. Acad. Sci. U.S.A.* 106, 10587–10592. doi: 10.1073/pnas.0903616106

PHONOLOGICAL GRAMMARS EVOLVE TO PRESERVE INFORMATION AT WORD BEGINNINGS

ANDREW WEDEL ^{*1}, ADAM USSISHKIN¹, ADAM KING¹, JAYCIE MARTIN¹, JOHN GEARY¹

*Corresponding Author: wedel@email.arizona.edu ¹Department of Linguistics, University of Arizona, Tucson, USA

1. Introduction

Listeners identify words incrementally at the sub-lexical level, continually updating hypotheses about the identity of the word as its phonetic signal unfolds. Because information is integrated incrementally, phonetic cues earlier in the word contribute more information on average to lexical identification than phonetic cues later in the word (e.g. Van Son & Pols 2003). For example, the initial [v] in the English word *vacuum* is highly informative, while the final [m] contributes less, because *vacuum* is already likely given the preceding signal [vækju-]. Recent work suggests that lexicons evolve to take advantage of this inherent bias, preferentially allocating more informative segments toward word-beginnings where they can contribute more to lexical access (King 2017; King & Wedel in prep; Meylan & Griffiths 2017). The phonetic form of words is not static however: all languages have a set of *phonological rules* which predictably modify pronunciation in context. As an example, all stop consonants in German are devoiced word-finally, such that 'Hund' *dog* is pronounced 'Hun[t]'.

Evidence suggests that phonological rules develop in a language when there are consistent biases on how speakers pronounce sounds in context. Because phonetic cues are on average most informative at word beginnings, and speakers pronounce informative phonetic cues more carefully than less informative cues (e.g., Aylett & Turk 2004; Wedel, Nelson & Sharp, to appear), we predict that languages should be less likely to evolve phonological rules which reduce lexical information at the beginnings of words. Here, we investigate this question through a statistical analysis of a cross-linguistic dataset of phonological rules.

2. Methods and Results

We assembled a genetically and areally-diverse dataset of phonological rules coding whether the rule modifies the beginning or end of a word. The dataset contains 266 rules from 50 languages (Figure 1). We dichotomously coded each rule in the dataset for whether the rule is phonemically neutralizing (i.e., whether it potentially creates homophones). Two patterns appear in this dataset: (i) there are significantly more rules overall that modify the ends of words, and (ii), this end-bias is significantly stronger for neutralizing rules (Figure 2). Both patterns are statistically significant in a mixed-effects regression model including Language, Family and Area as random intercepts. We further investigated two potential confounds that could explain this data: (i) a wellknown bias toward rules that modify syllable-final consonants, and (ii) the fact that suffixing-dominant languages are more common; given that phonological rules often arise at stem-affix boundaries, this could potentially account for an apparent end-bias in a sample of languages. However, we show that neither of these alternative explanations can account for the apparent bias for contrastreducing rules to target the ends of words.

These results provide the first statistical evidence supporting the hypothesis that languages evolve phonological grammars which preferentially preserve initial lexical material, which tends to be most informative in communication. More broadly, this finding contributes to the increasingly sophisticated body of evidence that language structures evolve under conflicting biases toward accurate transmission of meaning and effort reduction (e.g., Zipf 1949; Lindblom 1990; Piantadosi et al. 2011; Wedel et al. 2013; Futrell et al. 2015).

Figure 1. Areal language distribution

Figure 2. Rule type by word edge





- Aylett, M. & Turk, A. (2004). The Smooth Signal Redundancy Hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech* 47: 31-56.
- Futrell, R., Mahowald, K., & Gibson, E. (2015). Large-scale evidence of dependency length minimization in 37 languages. *Proceedings of the National Academy of Sciences*, 112: 10336-41.
- King, A. (2017). Word Informativity Affects Segmental Informativity. Unpublished Masters Thesis. University of Arizona, USA.
- King, A. & Wedel, A. (in prep). Greater word informativity predicts greater word-initial segment informativity.
- Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. *Speech production and speech modelling*, ed. by W.J. Hardcastle & A. Marchal, 403-39. Dordrecht: Kluwer.
- Meylan, S. C. & Griffiths, T. L. (2017). Word forms not just their lengths are optimized for efficient communication. Unpublished manuscript.
- Wedel, A., Nelson, N.R., & Sharp R. (to appear). The phonetic specificity of contrastive hyperarticulation in natural speech. *Journal of Memory and Language*.
- Piantadosi, S. T, Tily H. & Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences* 108: 3526-29.
- Van Son, R. J. J. H. & Louis C. W. Pols. (2003). How efficient is speech? Proceedings of the Institute of Phonetic Sciences 25: 171-84.
- Wedel, A., Kaplan, A. & Jackson, J. (2013). High functional load inhibits phonological contrast loss: A corpus study. *Cognition* 128: 179-86.
- Zipf, G. K. (1949). Human behavior and the principle of least effort: An introduction to human ecology. Addison-Wesley.

FROM CONTEXT TO CODE: SHARED INFORMATION CONSTRAINS THE EVOLUTION OF OPTIMAL CODES

JAMES WINTERS^{*1} and OLIVIER MORIN¹

^{*}Corresponding Author: winters@shh.mpg.de ¹The Minds and Traditions Research Group, Max Planck for SHH, Jena, Germany

We communicate by means of coded information, contained in spoken words, written signs, or gestures. A code is a set of reliable associations between signals and meanings used to facilitate information transfer (Scott-Phillips, 2014). The amount of information that can be extracted from a code depends on its stability (the regularity of a signal-meaning mapping), and on whether the set of signal-meaning associations form an optimal configuration. A code is optimal when it reaches a tradeoff between two constraints. First, it contains the relevant information that the Recipient is supposed to extract from it (*informativeness*). Second, it presents this information in a compressible format, using mappings with a minimal description length (*compressibility*). Optimal codes are therefore the most compressible set of signal-meaning mappings capable of identifying the intended meaning in context (Kirby et al., 2015).

Optimal codes help facilitate successful communication, but are not indispensable. Efficient communication can happen even with an unstable and weak code – because it makes use of information that is contextual as opposed to encoded (Piantadosi et al., 2012). Context, in this sense, is information external to a code which is relevant for reducing uncertainty about a Messenger's intended meaning. This contextual information can be retrieved from the common environment that communicators share; it can be stored in memory; or it can be inferred from a combination of these sources (Clark & Brennan, 1991). Only when the context is sufficiently rich to allow for some information transfer to take place, do optimal codes emerge to fill and enrich expressive gaps in communication (Bybee, 2010).

We test the hypothesis that the informational distance between a Messenger and a Recipient (i.e., the amount of *shared information*) acts as a

strong constraint on solutions to this context-coding problem. Affecting the amount of shared information is whether the mode of information transfer is *synchronous* or *asynchronous* and whether the function of information transfer is *monadic* or *dyadic*. Synchrony, which we define as information transfer within the same timeframe, leverages the presence of shared contextual information, as well as the ability to provide feedback about informativeness, to increase the amount of shared information. Asynchrony, on the other hand, is where information transfer takes place across timeframes, reducing access to shared contextual information and removing availability of immediate feedback. Monadic information transfer, where information is transferred within an individual, corresponds to an increase in shared information: individuals must only coordinate with themselves to facilitate successful information transfer. This is in contrast to when information transfer is dyadic: Messenger and Recipient are different individuals, making coordination a more substantial problem as there is less shared information.

Using a referential game set up, where participants play as Messengers and Recipients, we experimentally manipulate both the mode (synchronous or asynchronous) and function of information transfer (monad or dyad) to generate four conditions: Recall (synchronous + monad), Mnemonic (asynchronous + monad), Dialogue (synchronous + dyad), and Correspondence (asynchronous + dyad). Only in Dialogue, where Messengers and Recipients are two different persons interacting within the same timeframe, do we consistently observe the emergence of stable and optimal codes. In the Mnemonic condition, where Messenger and Recipient are the same person at different points in time, participants produce informative codes that are comparatively less stable and compressible than those in Dialogue. By contrast, in the Recall condition, where Messenger and Recipient are the same person at the same point in time, and the Correspondence condition, where Messenger and Recipient are the same person at the same point in time, and the messenger and Recipient are the same person at the same point in time, and the messenger and Recipient are the same person at the same point in time, and the messenger and Recipient are the same person at the same point in time, and the messenger and Recipient are the same person at the same point in time, and the messenger and Recipient are the same person at the same point in time, and the messenger and Recipient are the same person at the same point in time, and the messenger and Recipient are the same person at the same point in time, and the messenger and Recipient are the same person at the same point in time, and the correspondence condition, where Messenger and Recipient are two different persons communicating across timeframes, stable and optimal codes fail to emerge.

These results demonstrate that solutions to the context-coding problem are contingent on the amount of shared information. If interlocutors share access to the same perceptual context, and have recourse to immediate feedback on their performance, then Messengers can leverage this shared information to rapidly coordinate with Recipients in constructing an optimal code. This builds on the idea that human communication is adapted to synchronous interactions between individuals (Levinson, 2006); codes can readily leave out information already provided by the context, and miscommunications can be repaired on the fly. When the distance between interlocutors increases, as is the case in the Correspondence condition, Messengers are unable to estimate, and therefore exploit, the information they share with the Recipients, resulting in unstable and nonoptimal codes.

References

Bybee, J. (2010). Language, usage and cognition. Cambridge University Press.

- Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. Resnick, L. B., M. John, S. Teasley, & D. (Eds.), *Perspectives on socially shared cognition* (pp. 13–1991). American Psychological Association.
- Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, 141, 87–102.
- Levinson, S. C. (2006). On the human interaction engine. In N. J. E. S. C. Levinson (Ed.), *Roots of human sociality: Culture, cognition and interaction* (p. 39-69). Oxford: Berg.
- Piantadosi, S. T., Tily, H., & Gibson, E. (2012). The communicative function of ambiguity in language. *Cognition*, 122(3), 280–291.
- Scott-Phillips, T. C. (2014). Speaking Our Minds. Macmillan Education UK.

A MODEL OF CULTURAL CO-EVOLUTION OF LANGUAGE AND PERSPECTIVE-TAKING

MARIEKE WOENSDREGT*1, SIMON KIRBY1, CHRIS R. CUMMINS1, and KENNY SMITH1

*Corresponding Author: mwoensdr@ed.ac.uk ¹Centre for Language Evolution, University of Edinburgh, UK

Language relies on *theory of mind*, as language users have to entertain and recognise communicative intentions (e.g. Moore, 2016; Scott-Philips, 2014). Theory of mind abilities in turn profit from language, as language provides a means for expressing mental states explicitly (e.g. Bar-On, 2013), and for transmitting one's understanding of minds to others (e.g. younger members of the population) (Heyes & Frith, 2014). Given this interdependence, it has been hypothesised that language and theory of mind have co-evolved (e.g. Malle, 2002).

We present an agent-based model to formalise this hypothesis, which builds on a model described in Woensdregt, Kirby, Cummins, and Smith (2016). This model focuses on the ability to take others' perspectives, which is part of (the development of) theory of mind (e.g. Apperly, 2011), and has a clear and concrete role to play in everyday referential signalling (e.g. Clark & Wilkes-Gibbs, 1986). In this model, the communicative behaviour of an agent is determined not just by their language but also by their perspective on the world. Each agent has a perspective which, in interaction with a given context, determines the salience of potential topics to them. Because all objects in the world are considered potential topics, learners have no other way of inferring the intended referent of an utterance than by gaining knowledge of the speaker's perspective and lexicon (i.e. crosssituational learning is not possible). However, neither a speaker's perspective nor their lexicon are directly observable. Learners thus have to infer both of these simultaneously, using Bayesian inference. Simulation results show that learners can solve this task by bootstrapping one from the other, but only if the speaker uses a language that is at least somewhat informative (Woensdregt et al., 2016).

Given this model of how an agent's perspective in combination with a context gives rise to a mental state, we can compare two different types of agents. *Literal agents* choose an utterance purely based on whether or not the signal is associated with the intended referent in their lexicon. A listener then interprets the utterance by inverting this speaker model (using the lexicon and perspective hypothesis they selected for this speaker after a learning phase). Listeners of this type thus use their perspective-taking ability for interpreting utterances, but literal speakers do

not do so for production. *Pragmatic agents*, in contrast, reason about each other's minds in both directions. A pragmatic speaker optimises their utterances based on how a perspective-taking listener would interpret them (using the *rational speech act* model, Goodman & Frank, 2016). Moreover, pragmatic speakers can in some cases capitalise on the context (and the assumption that the listener knows their perspective) in order to mitigate the adverse effect of ambiguity in their lexicon.

We embedded these models of literal and pragmatic agents in an iterated learning model (where languages are transmitted over generations), in order to explore under which circumstances populations can build a fully informative language from scratch. With literal agents, we found that without any additional learning bias or selection pressure, populations do not establish any linguistic conventions. If we add selection on communicative success, however, populations do evolve a fully informative language. Interestingly, selection on correctly inferring perspectives also results in partially-informative lexicons emerging, which is sufficient for inferring others' perspectives. Note that the only thing that is transmitted over generations is the language; this is thus a model of cultural evolution, where better perspective-inference is reached only by virtue of the population establishing meaningful linguistic conventions. Selecting for the ability to infer perspectives introduces a pressure which leads to the cultural evolution of languages which permit such inference, i.e. the more informative ones. These results illustrate the potential for cultural co-evolution between language and perspective-taking, given the assumption that these two skills are interdependent in their development.

With pragmatic agents, we find the same result when no selection pressure is present. However, the results under selection look different, such that both pressures (selection for communication and selection for inferring perspectives) lead to similar levels of informativeness, and hence similar (high) levels of success at both communicating and inferring perspectives. This is because pragmatic agents can reach high levels of communicative success even with suboptimal languages, while both communication and inferring perspectives still rely on the language being somewhat informative.

To recap, the model described here assumes that language learning and perspective-inference develop interdependently. This is the same for both literal and pragmatic agents, but pragmatic agents do an extra step of reasoning about each others' minds when they communicate. This yields a cultural evolution model in which theory of mind is useful for communication, while language in turn helps to learn about others' minds. We show that when this is the case, a pressure to evolve one of these two skills can give rise to the other for free. In the case of literal agents, the level of communicative success that is reached under a selection pressure for the ability to infer perspectives is limited. However, in the case of pragmatic agents, both types of selection lead to equally high success in communication and perspective-inference, because these agents can communicate successfully even with languages that contain ambiguity.

- Apperly, I. (2011). *Mindreaders: The Cognitive Basis of "Theory of Mind"*. Psychology Press.
- Bar-On, D. (2013). Origins of Meaning: Must We 'Go Gricean'? Mind & Language, 28(3), 342–375.
- Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, 22, 1–39.
- Goodman, N. D., & Frank, M. C. (2016). Pragmatic Language Interpretation as Probabilistic Inference. *Trends in Cognitive Sciences*, 20(11), 818–829.
- Heyes, C. M., & Frith, C. D. (2014). The cultural evolution of mind reading. *Science*, *344*(6190), 1243091–1–6.
- Malle, B. F. (2002). The relation between language and theory of mind in development and evolution. In T. Givón & B. F. Malle (Eds.), *The evolution of language out of pre-language* (pp. 265–284). Amsterdam: John Benjamins.
- Moore, R. (2016). Gricean Communication and Cognitive Development. *The Philosophical Quarterly*, 67(267), 303–326.
- Scott-Philips, T. (2014). Speaking our minds. Palgrave Macmillan.
- Woensdregt, M. S., Kirby, S., Cummins, C., & Smith, K. (2016). Modelling the co-development of word learning and perspective-taking. *Proceedings of* the 38th Annual Meeting of the Cognitive Science Society, 1241–1246.

INFANTS' SENSITIVITY TO EMOTIONAL ANIMAL VOCALIZATION AND THE EVOLUTION OF VOCAL COMMUNICATION

NAOTO YAMANE^{*1}, MIHOKO HASEGAWA¹, AI KANATO^{1,2}, NAOKO KIJIMA¹, KAZUO OKANOYA^{3,} and REIKO MAZUKA^{1, 4}

*Corresponding Author: *yamanen@brain.riken.jp* ¹Laboratory for language development, RIKEN BSI, Saitama, Japan ²Waseda University, Saitama, Japan ³University of Tokyo, Tokyo, Japan ⁴Department of Psychology and Neuroscience, Duke University, Durham NC, USA

For many species, transmitting emotional signals is a fundamental function of vocal communication. In pre-verbal human infants, emotions conveyed by the speech signal are perhaps the first meanings that they are able to detect (Fernald, 1993). Indeed, human newborns were shown to present an increase in eye opening responses following the presentation of happy mother's voice as compared to the other emotions (Maestropieri & Turkewitz, 1999). By 5 months, human infants can respond appropriately to emotions conveyed by speech prosody (Fernald, 1993). Although the characteristics of vocalization that animal produce are often species specific, it appears that there is a general relationship between the physical structure of sounds and the motivation underlying their use (Morton, 1977). If different species share the ability to produce common emotional vocal signals, they may also share a common ability to detect such cues. To date, however, no studies have empirically examined whether humans or other species possess such an ability. In the present study, we examined whether human infants have sensitivities in emotional vocalizations of other species. In each of the 3 experiments, 5- and 9-month-old Japanese infants (20 in each age group) were tested using the Head-Turn Preference paradigm (HPP). The vocalizations of a naked mole rat (Exp. 1), a song sparrow (Exp. 2) and a swamp sparrow (Exp. 3) were chosen and were

This paper is distributed under a Creative Commons BY-NC-SA license. DOI:10.12775/3991-1.134

534

produced in either a positive (pleasure, contact) or a negative (distress, alarm) context. Infants' average listening times to the positive and the negative contexts were measured for each type of vocalizations. In Figure 1, the First Look listening time is the average listening time when the infant first oriented to the stimuli, and the Total Look listening time is the average of the total time during which infants oriented to the stimuli. For the vocalizations of the naked mole rat, infants in both age groups showed longer listening time to the negative vocalizations than the positive one both in the First Look [F(1, 38) = 120.56, p]< .01] and the Total Look [F (1, 38) = 161.37, p < .01]. For the song sparrow vocalizations, 5- and 9-month-olds listened longer to the positive vocalizations than the negative ones [First Look, F(1, 38) = 24.7, p < .01; Total Look, F(1, 38)= 30.5, p < .01]. For the swamp sparrow vocalizations, 5-month-olds but not 9month-olds listened longer to the negative ones in the First Look [F (1, 19) = 2.39, P < .05], while no difference for the Total Look [F (1, 38) = 2.26, n.s.] were found. These results could not be explained only by specific acoustic cues, such as pitch, loudness and duration. Hence, the results of our experiments clearly demonstrate that human infants possess an ability to detect some kinds of emotional vocal signals in other species. This indicates that human and nonhuman species may indeed possess a shared ability not only to produce common emotional vocal signals but also to perceive such cues.



Figure 1. Mean looking times (with standard error bars) of First Look (upper part) and Total Look (lower part) for the positive (white bars) and negative (black bars) vocalizations produced by Naked mole rat (left), Song sparrow (center) and Swamp sparrow (right).

Acknowledgements

This study was supported in part by Grant-in-Aid for Young Scientists (B) (JP16K16629) to NY, Scientific Research S (16H06319) and Scientific Research on Innovative Areas #4903 (Co-creative Language Evolution) (17H06382) to RM.

- Fernald, A. (1993). Approval and disapproval: Infant responsiveness to vocal affect in familiar and unfamiliar languages. *Child development*, 64(3), 657-674.
- Mastropieri, D., & Turkewitz, G. (1999). Prenatal experience and neonatal responsiveness to vocal expressions of emotion. *Developmental psychobiology*, 35(3), 204-214.
- Morton, E. S. (1977). On the occurrence and significance of motivationstructural rules in some bird and mammal sounds. *The American Naturalist*, 111(981), 855-869.

THE ENGLISH DATIVE ALTERNATION AS AN ADAPTATION TO CHANGES IN THE CONSTRUCTIONAL NETWORK

EVA ZEHENTNER*1

*Corresponding author: eva.zehentner@york.ac.uk ¹Department of Language and Linguistic Science, University of York, York, UK

This paper investigates the history of the English dative alternation, i.e. the fact that most ditransitive, three-participant verbs in Present Day English mainly occur in two different constructions, the double object construction (DOC, 1) on the one hand, and a prepositional construction with to (to-POC, 2) on the other hand.

- (1) John gave Mary <u>a book</u>.
- (2) John gave <u>a book</u> to Mary.

More specifically, the paper discusses the emergence of this alternation as an adaptive response to or evolutionary effect of changes in the constructional environment of the patterns involved. This is argued to work in a two-fold way. On the one side, the alternation as such is claimed to reflect adaptive changes to system-wide changes such as the loss of case marking or the increasing fixation of word order in the history of English. On the other side, the specific features of the members of the dative alternation are seen as the result of two constructions adapting to each other once they become linked in the network. Both these assumptions rest on and were tested on the basis of different methodologies:

First, an evolutionary game theoretic model (e.g. Hofbauer & Sigmund, 1998) was applied to the issue at hand. This method, originally a branch of applied mathematics, is used to "stud[y] the general problem of strategy selection and its propagation across a population" (Deo, 2015) and has recently also been extended to linguistics (e.g. Jaeger, 2008; Deo, 2015). The results of our game demonstrate that under certain conditions (reflecting universal principles such as end-focus) broader changes can indeed lead to the establishment of a close link between originally unrelated constructions.

538

Second, the findings of a large-scale corpus study of the Penn-Helsinki Parsed Corpus of Middle English (PPCME2) were drawn on to investigate the constructions' development in more detail. The data show that Middle English saw the resident DOC and the innovative to-POC entering into a state of stable co-existence, with the nominal construction as the stronger part, and the prepositional construction as the weaker one. This development is then taken to constitute the outcome of competition between two constructional variants, which has not resulted in the ousting of one competitor, but has instead led to the establishment of a cooperative relationship between the patterns (cf. e.g. Berg, 2014). That this association is mutually beneficial is supported by the existence of positive priming effects between the constructions (cf. Perek, 2015), but also by the division of labour-situation evidenced by them: Both patterns have differentiated according to discourse-pragmatic features such as givenness (e.g. Bresnan et al., 2007). Interestingly enough, however, the constructions also exhibit signs of what has recently been dubbed 'attraction' (De Smet et al., subm.), meaning that the variants have formally and functionally aligned to each other. The constructions have thus differentiated in some aspects, but have become more similar in respect to others. In this paper, both niche construction/ differentation and attraction are interpreted as indicators of constructional 'coevolution', i.e. constructions mutually adapting to each other.

- Berg, T. (2014). Competition as a unifying concept for the study of language. *The Mental Lexicon*, *9*, 338-370.
- Bresnan, J., Cueni, A., Nikitina, T. & Baayen, R.H. (2007). Predicting the dative alternation. In G. Bouma, I. Kraemer and J. Zwarts (Eds.), *Cognitive foundations of interpretation* (pp. 64-94). Amsterdam: Royal Netherlands Academy of Science.
- Deo, A. (2015). The semantic and pragmatic underpinnings of grammaticalization paths: The progressive and the imperfective. *Semantics and Pragmatics*, 8(14), 1-52.
- De Smet, H., D'Hoedt, F., Fonteyn, L. & van Goethem, C. (submitted). The changing functions of competing forms: Attraction and differentiation.
- Hofbauer, J. & Sigmund, K. (1998). *Evolutionary games and population dynamics*. Cambridge: Cambridge University Press.
- Jaeger, G. (2008). Applications of game theory in linguistics. Language and Linguistics Compass, 2(3), 408-421.
- Perek, F. (2015). Argument structure in usage-based construction grammar: Experimental and corpus-based perspectives. Amsterdam: Benjamins.

USING LANGUAGE TO DISCOVER CATEGORIES: MORE NAMEABLE FEATURES IMPROVE CATEGORY LEARNING

MARTIN ZETTERSTEN*1 and GARY LUPYAN1

*Corresponding Author: zettersten@wisc.edu ¹Psychology Department, University of Wisconsin-Madison, Madison, USA

Does the lexicon of a language have consequences for cognition? Here, we provide evidence that the ease with which category features can be named can influence category learning. Across two experiments, participants learned to distinguish images composed of colors (Experiment 1) and shapes (Experiment 2) that were either easy or more difficult to name in English. Holding the category structure constant, when the underlying features of the category were easy to name, participants were faster and more accurate in learning the novel category. We argue that these findings suggest that labels allow learners to form more compact hypotheses, which in turn can be confirmed or disconfirmed in the course of learning. These results have consequences for considering how cross-linguistic differences in lexical inventory affect how readily novel categories are learned.

1. Introduction

Categorization is central to cognition (Harnad, 1990; Murphy, 2002). Learning categories requires recognizing which features or dimensions of stimuli are relevant and which are not. For example, color is a relevant dimension for categorizing foods because many foods have characteristic colors. In contrast, color is largely irrelevant for categorizing vehicles: knowing that an object is brown does not help in categorizing it as a car. But how do we discover which features and dimensions are relevant to category membership? One common proposal is that learners track how strongly different features are associated with a category (Rosch & Mervis, 1975). But how do learners identify candidate features in the first place (see e.g., Schyns, Goldstone, & Thibaut, 1998 for discussion)?

Here, we ask whether part of the answer lies in language (Lupyan, 2012). Might the vocabulary of a language provide the learner with a powerful set of candidate features (priors) to use when learning new categories? In particular, we test the hypothesis that the ease with which features can be named influences the likelihood that they are considered as relevant for category membership. For example, on encountering a novel stimulus, learners can more easily form a verbal description of some candidate features (e.g., "red" as a description of a color or "tree" as a description of a complex shape) compared to others (e.g., "greenish-yellow-brown mix" or "nonsymmetrical spacecraft looking thing"). Features that are more nameable may be represented in a more stable way,

making it more likely to provide a hypothesis about category membership (e.g., "items with the color red belong in the same category") that can be tested during category learning. If language plays a role in our ability to form hypotheses about novel categories, this has important implications for understanding the consequences of cross-linguistic differences: when a language has a readily accessible name for a feature or dimension, it makes that feature/dimension easier to represent during category learning.

In the current experiment, we compare participants' ability to learn novel categories when category exemplars were composed of more nameable or less nameable color features (Experiment 1) or shape features (Experiment 2). Both category types had similar logical structure: the categories were structured such that one color feature or one shape feature always predicted category membership (e.g., images containing red belonged to category A while those containing brown belonged to category B). We hypothesized that categories were easier to learn when the underlying color and shape features were easier to name, and therefore could be more readily formulated as a hypothesis about category membership.

2. Experiment 1: Color-based category features

2.1. Participants

We recruited 201 participants through Amazon Mechanical Turk. Participants were randomly assigned to the High Nameability Condition (n = 101) or to the Low Nameability Condition (n = 100) and were paid \$0.60 - \$0.75 for completing the task, which lasted approximately 4 minutes.

2.2. Stimuli

The exemplars were circles ("color wheels") composed of 3 different colors (see Figure 1A). Following the design of (Couchman, Coutinho, & Smith, 2010), one of the colors was perfectly predictive of category membership. The other two were correlated at 66.6% with category membership. The critical manipulation involved the nameability (Guest & Laar, 2002) of the colors comprising each color wheel exemplar. To assess nameability we used the results of a large-scale online color naming study (N=134,727, Munroe, 2010), restricting the analysis to RGB values named by \geq 100 individuals. To ensure that the results could not be affected by differences in color discriminability between the low- and high-nameability conditions, we selected the colors such that all the pairwise CIE-LAB distances were equated using Δ E2000 (Sharma et al., 2005). We

constructed image sets with different color pairs in the critical position to ensure that any effect was not due to idiosyncratic properties of colors in the categorydiagnostic position. The critical colors in the more-nameable categories were "brown" RGB=(120, 80, 40) and "red" RGB=(220, 20, 0), "blue" RGB=(30, 90, 210) and "orange" RGB=(250, 120, 30), or "blue" and "brown". These colors were named according to their modal label by 80% - 85% of the population. The critical colors in the low-nameability were RGB=(170,160,40) (modal name "mustard") and RGB=(200,170,170) (modal name "lavender") or RGB=(200, 100, 70) (modal name "brown") and RGB=(70, 100, 90) (modal name "grey"). These names were only used by 6% - 10% of participants in the original naming task. One drawback of the current approach is that ΔE requires a device-independent color space while our participants viewed the colors on their (often uncalibrated) monitors. However, while this introduces noise, there is no indication that this variability in viewing conditions favorably affects the high-nameability colors compared to the low-nameability colors.

2.3. Task

The participants were tasked with learning to place the color wheels into one of two categories by dragging it into the appropriate box , labeled as box "A" and box "B" (Figure 2). Participants completed a total of 24 training trials, split into 3 blocks. On each block, participants sorted the prototype exemplar (the top image in Figure 1A) for each category twice, and the remaining two exemplars of each category once. Participants received immediate feedback on whether their choice was correct or incorrect. Trials were repeated after an incorrect response. Box locations were counterbalanced across participants.



Figure 1. (A) Example stimuli for the high and low nameability condition and (B) Accuracy on categorization across blocks depending on nameability condition in Experiment 1 (color-based features).



Figure 2. Task setup for category learning experiment

2.4. Results

Training results are shown in Figure 1B. We tested the effects of condition (low nameability vs. high nameability), block, and their interaction on participants' accuracy using a logistic mixed-effects model. Variables were centered and the model included a by-subject random intercept and random slope for block.

Accuracy increased over the course of training, b = .98, Wald 95% CI = [.81, 1.16], z = 11.03, p < .0001. Overall accuracy was higher in the high nameability condition (M = 85.1%, 95% CI = [82.7%, 87.5%]) compared to the low nameability condition (M = 77.4%, 95% CI = [74.3%, 80.4%]), b = 0.74, Wald 95% CI = [.39, 1.09], z = 4.13, p < .0001. There was significant interaction between block and condition, b = .43, Wald 95% CI = [.14, .71], z = 2.89, p = .003, indicating that participants learned the categories more quickly in the high nameability condition than in the low nameability condition.

3. Experiment 2: Shape-based category features

3.1. Participants

We recruited 120 participants through Amazon Mechanical Turk. Participants were randomly assigned to the High Nameability Condition (n = 58) or to the Low Nameability Condition (n = 62) and were paid \$0.90 for completing the task, which lasted 8 minutes on average.

3.2. Stimuli

The exemplars were circles similar to Experiment 1 composed of 2 different shapes (see Figure 3A). The shapes were chosen from a previous study on the nameability of complex polygon shapes (Vanderplas & Garvin, 1959). To ensure that shapes used in the experiment were equally discriminable, we conducted a separate norming task in which participants (n = 24) performed a speeded same/different task—an extremely sensitive method for measuring represented visual similarity (Lupyan, 2008). We collected these visual discriminability data for the eight most nameable and the eight least nameable shapes from Vanderplas & Garvin (1959)'s ratings. The high nameability shapes were slightly more discriminable than the low nameability shapes, making it impossible to find enough shapes to recreate the same feature structure as in Experiment 1 while still matching the two conditions on shape discriminability. We therefore selected four highly nameable shapes and four shapes with low nameability such that the pairwise discriminability (as measured by reaction times in the norming task) was matched between the two shape sets. The modal

names for the highly nameable shapes were "bear", "tree", "frog", and "swan" (left to right in Fig. 3A). The modal names for the low nameability shapes were "mountain", "rabbit", "bird" and "hook". As in Experiment 1, the two categories were defined by the presence of a single critical shape. To increase the difficulty

of the task and increase the variability of the stimuli, we randomized the location of the shapes, such that the critical image appeared in all three "slice" locations rather than in a fixed location, as in Experiment 1.

3.3. Task

The category learning task was identical to Experiment 1. Participants completed 24 training trials split into 3 blocks.



Figure 3. (A) Example stimuli for the high and low nameability condition and (B) Accuracy on categorization across blocks depending on nameability condition in Experiment 2 (shape-based features).

3.4. Results

To analyze the training results (see Figure 3B), we fit the same model as in Experiment 1. Accuracy increased over training blocks, b = 1.45, Wald 95% CI = [1.17, 1.73], z = 10.09, p < .0001. As in Experiment 1, overall accuracy was higher in the high nameability condition (M = 81.5%, 95% CI = [77.7%,

85.6%]) compared to the low nameability condition (M = 66.3%, 95% CI = [61.7%, 70.8%]), b = 1.45, Wald 95% CI = [.89, 2.01], z = 5.04, p < .0001. We also found a significant interaction between block and condition, b = .91, Wald 95% CI = [.42, 1.39], z = 3.67, p < .001, indicating faster increases in accuracy in the high nameability condition compared to the low nameability condition.

4. Discussion

Categories defined by a more nameable color or shape were learned substantially faster than categories with an identical structure but whose defining feature was less nameable. This result supports the idea that, at least for learning simple categories with diagnostic features, more easily formed verbal descriptions may make it easier for participants to formulate and test hypotheses about which features define a category. This result supports and extends previous work showing that controlling for categories (Lupyan, Rakison, & McClelland, 2007).

An important question is whether labels truly are the causal force driving the difference in categorization accuracy. One potential worry is that more nameable colors and shapes are more nameable *because* they are somehow cognitively simpler, rather than because labels aid learners in representing them in the context of category learning. One way to address this worry is to teach labels for the low-nameability categories and test the specific efficacy of such label-based training. Another way is to conduct cross-linguistic experiments: to the extent that language does the driving, categorization performance of e.g., Mandarin speakers, ought to be better predicted by shape/color nameability of Mandarin rather than English. We are currently using both approaches to address the question of causality.

Another worry is that nameability is confounded with familiarity. Perhaps more nameable items are more familiar and it is familiarity rather than language that is causing the difference between conditions. This is possible in the case of shapes (though note that the *specific* shapes were equally novel). We are unsure how a familiarity-based explanation would work in the case of color, however. Unsaturated colors are considerably harder to name than saturated colors. However, saturated colors span only a small region of color space, and most of our experiences involve less than fully saturated colors, making less saturated colors, if anything, *more* perceptually familiar.

The current findings have important implications for thinking about why and in what way cross-linguistic differences in vocabulary may impact cognition. When one language provides a compact label for a feature that another language does not, it provides a verbal encoding that allows learners to more readily represent this feature. A more easily represented feature can then be entertained as a potential high-level hypothesis about the structure of a novel category. This may help to explain the role played by language in performance across various domains, including representing exact numerosity (Frank, Everett, Fedorenko, & Gibson, 2008; Gordon, 2004) and relational reasoning (Christie & Gentner, 2014; Gentner, Ozyürek, Gürcanli, & Goldin-Meadow, 2013). As languages develop, they also change the verbal repertoire available to their users. These changes in verbal repertoire not only have consequences for communication, they also have consequences for forming novel categories.

Acknowledgements

This material is based upon work supported by NSF-PAC 1331293 to GL and NSF-GRFP DGE-1256259 to MZ. We thank Lynn K. Perry for developing central ideas underlying this study.

- Christie, S., & Gentner, D. (2014). Language helps children succeed on a classic analogy task. *Cognitive Science*, *38*(2), 383–397.
- Couchman, J. J., Coutinho, M. V. C., & Smith, J. D. (2010). Rules and resemblance: Their changing balance in the category learning of humans (Homo sapiens) and monkeys (macaca mulatta). *Journal of Experimental Psychology. Animal Behavior Processes*, 36(2), 172–183.
- Frank, M. C., Everett, D. L., Fedorenko, E., & Gibson, E. (2008). Number as a cognitive technology: Evidence from Pirahã language and cognition. *Cognition*, 108(3), 819–24.
- Gentner, D., Ozyürek, A., Gürcanli, O., & Goldin-Meadow, S. (2013). Spatial language facilitates spatial cognition: Evidence from children who lack language input. *Cognition*, 127(3), 318–30.
- Gordon, P. (2004). Numerical cognition without words: Evidence from Amazonia. *Science*, *306*(5695), 496–9.
- Guest, S., & Laar, D. V. (2002). The effect of name category and discriminability on the search characteristics of colour sets. *Perception*, 31(4), 445–461.
- Harnad, S. (Ed.). (1990). Categorical perception: The groundwork of cognition. Cambridge: Cambridge University Press.
- Lupyan, G. (2008). The conceptual grouping effect: Categories matter (and named categories matter more). *Cognition*, *108*(2), 566–577.

- Lupyan, G. (2012). What do words do? Toward a theory of language-augmented thought. In Brian H. Ross (Ed.), *Psychology of Learning and Motivation* (Volume 57, pp. 255–297). Academic Press.
- Lupyan, G., Rakison, D. H., & McClelland, J. L. (2007). Language is not just for talking: Labels facilitate learning of novel categories. *Psychological Science*, 18(12), 1077–1082.
- Munroe, R. P. (2010, May 4). Color survey results. Retrieved January 30, 2016, from http://blog.xkcd.com/2010/05/03/color-survey-results/
- Murphy, G. L. (2002). The big book of concepts. Cambridge, MA: MIT Press.
- Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7(4), 573–605.
- Schyns, P. G., Goldstone, R. L., & Thibaut, J. P. (1998). The development of features in object concepts. *The Behavioral and Brain Sciences*, 21(1), 1-17-54.
- Sharma, G., Wu, W. C., & Daa, E. N. (2005). The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. *Color Research and Application*, 30(1), 21–30.
- Vanderplas, J. M., & Garvin, E. A. (1959). The association value of random shapes. *Journal of Experimental Psychology*, 57(3), 147–154.

PANTOLANG: A SYNTHETIC COGNITIVE-SEMIOTIC APPROACH TO LANGUAGE ORIGINS

JORDAN ZLATEV^{*1}, INES ADORNETTI³, ALLESANDRA CHIERA³, SIMON DEVYLDER¹, FRANCESCO FERRETTI³, SLAWOMIR WACEWICZ², PRZEMYSŁAW ŻYWICZYŃSKI²

*Corresponding Author: jordan.zlatev@semiotik.lu.se
¹Centre for Languages and Literature, Lund University, Sweden
²Centre for Language Evolution Studies, Nicolaus Copernicus University, Torun, Poland
³Department of Philosophy, Roma Tre University, Rome, Italy

1. Introduction

We present an ongoing international project, *From Pantomime to Language* (PANTOLANG), aiming to develop a comprehensive, empirically grounded theory of the evolution of human language and the human mind, relying on the new paradigm of *cognitive semiotics*, which combines methods and concepts from the humanities and the sciences (Zlatev, 2015; Zlatev, et al. 2016).

2. Main concepts

The key concept of the project is that of *pantomime*, a communication system based on whole-body re-enactment of events, relying predominantly on iconicity/resemblance (Zywiczynski et al., 2016; Zlatev et al., 2017). As the foremost communicative manifestation of the uniquely human capacity for *bodily mimesis* (Donald, 2001; Zlatev, 2014), pantomime arguably introduced a new level of semiotic complexity: an open system of *signs*, rather than a closed system of association-based *signals*. While other theories have appealed to "gesture" or even "pantomime" as a precursor to language (e.g. Arbib, 2005; Tomasello, 2008), our approach is unique in defining the notion consistently and making it the cornerstone of a theory of language origins.

Further, to explain the transition from pantomime to language, we focus on three central cognitive-semiotic factors. The first is *intersubjectivity*, which implies human-specific levels of (mind) sharing and trust. We distinguish

between primary forms such as *emotional empathy* which were prerequisites for the emergence of pantomime, and secondary forms of intersubjectivity which evolved along with it, such as the following four features of human *interactivity*: alternation of turns, synchrony, conditional relevance and role-reversal (Sacks & Schegloff, 1973; Wacewicz & Zywiczynski, in press)

The second factor is that of *narrativity*, implying temporal and causal coherence between events, on the one hand, and their representations (narrations) on the other. As with intersubjectivity, we hold that there were precommunicative aspects of narrativity, allowing ancient hominins to perceive, understand and remember event-sequences or episodes, and assist them in navigation (Ferretti et al., 2016, 2017). But it is only pantomime that made externalization possible, so that hominins were first able to re-enact, and thus embody narrative structures in a publicly accessible manner. Present-day elaborated versions of such embodied narratives may be seen in the performances of traditional societies (e.g. Green, 2014).

The third factor is *cross-modality*, understood as the interaction between different sensory channels (rather than "communicative modes"). While pantomime was primarily perceived visually, it was also "felt" via the cross-modal mapping between vision and proprioception (Zlatev, 2014). While vocalizations are unlikely as the initial channel for referential meaning (Zlatev et al. 2017), they would have become increasingly important over a prolonged period of time – but without ever fully displacing pantomime, which appears in reduced form as iconic gestures in spontaneous language use. This distinguishes our approach from those who assume that language was "multimodal" from the onset (e.g. McNeill, 2012).

3. Methods

Following the principle of methodological triangulation of cognitive semiotics, combining philosophy, (participant) observation and experiments, we (a) use phenomenology and conceptual analysis to propose clear definitions of central concepts (e.g. Zlatev, 2015), (b) study polysemiotic narratives in traditional societies, such as Paamese "sand drawings" in Vanuatu, where pantomime, speech and other semiotic systems such as depiction combine (Devylder, 2014), and (c) adapt experimental semiotics paradigms (Galantucci & Garrod, 2011) to study the communicative effectiveness of pantomime across different cultures, and its conventionalization and communicative "streamlining" through horizontal (Fay et al, 2010) and vertical (Kirby et al., 2014) transmission.

- Arbib, M. (2005). From monkey-like action recognition to human language: an evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences*, 28, 105-168.
- Galantucci B & Garrod S. (2011). Experimental semiotics: a review. *Frontiers* in Human Neuroscience 5(11).
- Green, J. (2014). Drawn from the ground: Sound, sign and inscription in Central Australian sand stories. Cambridge: Cambridge University Press.
- Devylder, S. (2014) Paamese Language and Culture (Publication no. DOI:10.4225/72/58ab0479d6eeb).PARADISEC
- Donald, M. (2001). *A mind so rare: The evolution of human consciousness*. New York: Norton.
- Fay N., Garrod S., Roberts L., Swoboda N. (2010). The interactive evolution of human communication systems. *Cognitive Science* 34, 351–386
- Ferretti, F. & Adornetti, I. (2016). Mindreading, mind-travelling, and the protodiscursive origins of language. In Zlatev, et al. (2016) (eds.).
- Ferretti, F. Adornetti, I., Chiera, A., Nicchiarelli, S., Magni R., Valeri G., & Marini, A. (2017). Mental time travel and language evolution: A narrative account of the origins of human communication. *Lang Sci* 63, 105–118.
- Kirby, S, Griffiths, T & Smith, K. (2014). Iterated learning and the evolution of language. *Current Opinion in Neurobiology*, 28,108–114.
- McNeill, D. (2012). *How language began: Gesture and speech in human evolution*. Cambridge: Cambridge University Press.
- Sacks, H., & Schegloff, E. (1973). Opening up closings. Semiotica, 7(4), 289–327.
- Tomasello, M. (2008). *The origins of human communication*. Cambridge, MA: MIT Press.
- Wacewicz, S., & Zywiczynski, P. (in press). Language origins from an ultimate perspective: cooperation, trust, and turn-taking. *Interaction Studies*.
- Zlatev, J. (2014). Bodily mimesis and the transition to speech. In M. Pina & N. Gontier (Eds.), *The evolution of social communication in primates* (pp. 165-178). Berlin: Springer.
- Zlatev, J. (2015). Cognitive semiotics. In Trifonas, P. (ed.), *International* handbook of semiotics, 1043–1067. Dordrecht: Springer.
- Zlatev, J., Wacewicz, S., Zywiczynski, P. & van de Weijer, J. (2017). Multimodal-first or pantomime-first? Communicating events through pantomime with and without vocalization. *Interaction Studies* 18:3, 455– 479.
- Zlatev, J., Sonesson, G., & Konderak, P. (eds.) (2016). *Meaning, mind and communication. Explorations in cognitive semiotics.* Frankfurt/M.: Peter Lang.
- Żywiczyński, P., Wacewicz, S., & Sibierska, M. (2016) Defining pantomime for language evolution research. *Topoi*. doi: 10.1007/s11245-016-9425-9.

Vocal accommodation in common marmosets: Does similarity buffer tension during pair bond development?

YVONNE ZÜRCHER*1, ERIK P. WILLEMS1, JUDITH M. BURKART1

*Corresponding Author: Yvonne.zuercher@uzh.ch ¹Department of Anthropology, University of Zürich, Zürich, Switzerland

Affiliative relationships in animals are often based on similarity which can facilitate understanding and prediction of the behavior of the social partner (Capitanio et al., 2017; Massen and Koski, 2014; Morton et al., 2015; Weinstein and Capitanio, 2008). Similarity can influence the preference for a partner also in humans, and facilitate cooperation within a dyad. For instance, similarity in vocal dialect seems to increase cooperativeness in humans (Cohen and Haun, 2013; Heblich et al., 2015; Richerson et al., 2016).

Vocal accommodation occurs when communication partners adjust their vocalization or speech to a current interaction partner and can provide information about the social relationship between partners. It can take the form of convergence, when communication partners with a strong social bond become more similar, and therefore indicate social closeness. Alternatively, it can take the form of divergence, when partners' vocalizations become less similar, indicating social distance (Ruch et al., 2017). Since both humans and nonhuman primates appear to preferably interact with others that are similar to oneself (i.e. show homophily, (Haun and Over, 2015)), vocal accommodation may enhance group cohesion and buffer negative effects of vocal difference. The phenomenon of vocal accommodation is well studied in humans but less is known about it in nonhuman primates.

Callithrichids are highly vocal monkey species, and different populations show colony specific vocal differences (de la Torre and Snowdon, 2009; Zürcher and Burkart, 2017). Like humans, but unlike more closely related great apes, callitrichids are cooperative breeders and mothers strongly rely on help form their partners and other group members to raise offspring (Snowdon and Cronin, 2007; Snowdon and Ziegler, 2007). In callitrichid common marmoset (*Callithrix jacchus*) breeding pairs, pair bond strength correlates with cooperation in infant care (Finkenwirth, 2016). Furthermore, common marmoset pairs and their offspring show considerable behavioral similarity (group-level personalities that are independent of genetic relatedness: (Koski and Burkart, 2015)). Together, this suggests that it may be favorable for marmosets to be similar to their partners. However, choosing a similar mate may often not be possible due to limited availability of potential partners. In these cases, the ability to *become* more similar to a new mate would be beneficial.

552

Our goal was to investigate whether acoustic similarity improves pair bond formation in common marmosets, and whether they engage in vocal accommodation to reduce vocal distance. First, we expected that individuals with more similar vocalizations would develop a stronger pair bond than animals with different vocal dialects. To test this, we formed 10 new breeding pairs, 5 with animals with the same and 5 with different vocal dialects (similar / mixed pairs respectively). We collected behavioural data to quantify bond quality, expressed as affiliation, over the time span of the first 9 weeks of pair formation. The results showed that affiliation in mixed pairs developed significantly different over time compared to similar pairs (LME, p = 0.001). Mixed pairs showed a decrease of affiliation over time, whereas affiliation in similar pairs stayed constantly high. Second, we expected that pairs would show vocal accommodation. We had two hypotheses about why common marmosets would accommodate to their partner: 1) to advertise the strength of their pair bond or, 2) as a mean to buffer potential negative effects of vocal dissimilarity on bond quality. In the first case, we expected a positive correlation between the amount of accommodation and pair bond, in the second, more accommodation the larger the initial vocal distance was. To investigate if common marmosets show accommodation, we collected vocalizations of all new breeding partners over the course of the pair formation process (up to 9 weeks after first introduction). We analyzed three social call types (phee calls, trill calls and food calls) and measured 16 - 18 parameters per call, depending on call type. To estimate similarity between partners, we calculated Euclidian distances between the calls of partners at different time points and compared the differences with randomized t-tests. We found accommodation both in the form of convergence and divergence, with convergence being more common. Furthermore, the results did not support hypothesis 1 because vocal accommodation did not reflect average or final pair bond strength of a pair. Rather, in support of hypothesis 2, pairs with a larger initial vocal distance showed more convergence than pairs with smaller initial distance to their partner (LME, p = 0.001). We therefore suggest that 1) vocal similarity at the beginning of pair formation influences how the pair bond develops and 2) that common marmosets indeed use long term accommodation as a strategy to become more similar to their partner, buffering potential negative effects of vocal distance between partners. Taken together, our results suggest that social functions of accommodation did not evolve after, but rather before the evolution of language, probably serving as a base upon which language could evolve.

References

Capitanio, J. P., Blozis, S. A., Snarr, J., Steward, A., & McCowan, B. J. (2017). Do "Birds of a Feather Flock Together" or Do "Opposites Attract"? Behavioral Responses and Temperament Predict Success in Pairings of Rhesus Monkeys in a Laboratory Setting. *Am J Primatol*, 79(1), 1-11.

- Cohen, E., & Haun, D. (2013). The Development of Tag-Based Cooperation Via a Socially Acquired Trait. *Evolution and Human Behavior*, 34(3), 230-235.
- de la Torre, S., & Snowdon, C. T. (2009). Dialects in Pygmy Marmosets? Population Variation in Call Structure. *American Journal of Primatology*, 71(4), 333-342.
- Finkenwirth, C. (2016). *Individual Variation in Prosociality of Group-Living Common Marmosets: Assessing the Role of Oxytocin and Social Bonding*. Unpublished Dissertation, University of Zurich, Zürich.
- Haun, D., & Over, H. (2015). Like Me: A Homophily-Based Account of Human Culture. In *Epistemological Dimensions of Evolutionary Psychology* (pp. 117-130): Springer.
- Heblich, S., Lameli, A., & Riener, G. (2015). The Effect of Perceived Regional Accents on Individual Economic Behavior: A Lab Experiment on Linguistic Performance, Cognitive Ratings and Economic Decisions. *PLoS One*, 10(2), e0113475.
- Koski, S. E., & Burkart, J. M. (2015). Common Marmosets Show Social Plasticity and Group-Level Similarity in Personality. *Scientific reports*, 5.
- Massen, J. J. M., & Koski, S. E. (2014). Chimps of a Feather Sit Together: Chimpanzee Friendships Are Based on Homophily in Personality. *Evolution* and Human Behavior, 35(1), 1-8.
- Morton, F. B., Weiss, A., Buchanan-Smith, H. M., & Lee, P. C. (2015). Capuchin Monkeys with Similar Personalities Have Higher-Quality Relationships Independent of Age, Sex, Kinship and Rank. *Animal Behaviour*, 105, 163-171.
- Richerson, P., Baldini, R., Bell, A. V., Demps, K., Frost, K., Hillis, V., et al. (2016). Cultural Group Selection Follows Darwin's Classic Syllogism for the Operation of Selection. *Behavioral and Brain Sciences*, 39.
- Ruch, H., Zürcher, Y., & Burkart, J. M. (2017). The Function and Mechanism of Vocal Accommodation in Humans and Other Primates. *Biological Reviews*.
- Snowdon, C. T., & Cronin, K. A. (2007). Cooperative Breeders Do Cooperate. Behaviour Processes, 76(2), 138 - 141.
- Snowdon, C. T., & Ziegler, T. E. (2007). Growing up Cooperatively: Family Processes and Infant Care in Marmosets and Tamarins. *Journal of Developmental Processes*, 2(1), 40-66.
- Weinstein, T. A., & Capitanio, J. P. (2008). Individual Differences in Infant Temperament Predict Social Relationships of Yearling Rhesus Monkeys (Macaca Mulatta). Anim Behav, 76(2), 455-465.
- Zürcher, Y., & Burkart, J. M. (2017). Evidence for Dialects in Three Captive Populations of Common Marmosets (Callithrix Jacchus). *International Journal of Primatology*, 38(4), 780-793.

Author Index

Abels, Monika, 8 Adornetti, Ines, 548 Allritz, Matthias, 278 Amenta, Simona, 83 Archetti, Marco, 376 Arita, Takaya, 165 Asano, Rie, 12, 507 Barrett, H. Clark, 472 Baumann, Andreas, 15, 19, 408 Behme, Christina, 31 Bentz, Chris, 122 Bentz, Christian, 38 Berdicevskis, Aleksandrs, 41, 439Berscia, Luis Miguel, 252 Beuls, Katrien, 379 Blythe, Richard A., 189 Boeckx, Cedric, 44, 284 Bohn, Manuel, 278 Boruta, Monika, 50 Broadway, Megan, 270 Burkart, Judith M., 551 Burlak, Svetlana, 53 Byun, Kang-Suk, 67 Carr, Jon W., 70 Cartmill, Erica, 73, 472 Celma-Miralles, Alexandre, 76

Ceolin, Andrea, 196

Chiera, Alessandra, 548 Cichon, Kamil, 206 Cook, Dan, 80 Cordoni, Guido, 196 Crepaldi, Davide, 83 Culbertson, Jennifer, 70, 212, 215, 326, 469 Cummins, Chris, 531 Cuskley, Christine, 86

de Backer, Charlotte, 431 de Boer, Bart, 47, 399 de Gregorio, Chiara, 134 de Reus, Koen, 399 de Vos, Connie, 67, 522 Declerck, Carolyn, 431 Dediu, Dan, 38, 171 Delgado, Tania, 89 Devylder, Simon, 548 Dockum, Rikker, 92 Dreszer, Joanna, 50 Duch, Włodzisław, 50

Ellison, T. Mark, 95, 98 Emmorey, Karen, 2 Engesser, Sabrina, 101 Ergin, Rabia, 104

Fay, Nicolas, 316 Feeney, Andrew, 107
Feher, Olga, 111 Ferdinand, Vanessa, 114 Ferrer-I-Cancho, Ramon, 122, 161Ferretti, Francesco, 548 Flaherty, Molly, 125 Frank, Stella, 128 Franke, Michael, 158 Friard, Olivier, 134 Fujita, Koji, 131 Fusaroli, Riccardo, 349 Gärdenfors, Peter, 137 Gamba, Marco, 134 García, Ana Rubio, 399 Garcia, Maxime, 399 Gauvrit, Nicolas, 206 Geary, Jonathon, 525 Giacoma, Cristina, 134 Gleitman, Lila, 104 Goldberg, Adele, 3 Goldin-Meadow, Susan, 125, 405, 464Gontier, Nathalie, 139 Goodman, Noah, 158 Green, Glenn, 150 Green, Kari, 150 Grieve, Jack, 153 Gross, Stephanie, 399 Grossi, Thomas, 89 Gruber, Thibaud, 461 Grzybowska, Monika, 50 Guardiano, Cristina, 196 Hasegawa, Mihoko, 534 Hashimoto, Takashi, 507 Hawkins, Robert, 158 Hayashi, Akiko, 290 Heesen, Raphaela, 161 Higashi, Masahiko, 165

Hobaiter, Cat, 467

Hobaiter, Catherine, 161 Hoeksema, Nienke, 399 Horton, Laura, 405 Ignaczewska, Agnieszka, 50 Imai, Hirohiko, 346 Irimia, Monica, 196 Irvine, Elizabeth, 168 Jäger, Gerhard, 38 Jackendoff, Ray, 104 Janssen, Rick, 171 Jarvinen, Hanna, 111 Johansson, Niklas, 175 Johansson, Sverker, 178 Kalociński, Dariusz, 181, 367 Kanato, Ai, 534 Karjus, Andres, 189 Kashima, Eri, 98 Katsu, Noriko, 193 Kazakov, Dimitar, 196 Kempe, Vera, 206 Kijima, Naoko, 534 Kim, Shin-Sook, 196 King, Adam, 209, 525 Kirby, Simon, 70, 89, 189, 212, 215, 326, 370, 425, 492, 531Kirton, Fiona, 215 Kmiecik, Magdalena, 50 Knight, Chris, 218 Kobayashi, Harumi, 230 Kondo, Tadahisa, 290 Kuhle, Anneliese, 232 Lahiff, Nicole, 467 Lattenkamp, Ella, 235, 516 Lavrichenko, Ksenia, 238, 516 Leavens, David, 241 Lev-Ari, Shiri, 402

Levinson, Stephen C., 67 Lewis, Jerome, 4 Lewis, Molly, 243 Little, Hannah, 246, 419 Longobardi, Giuseppe, 196 Lou-Magnuson, Matthew, 249 Lupyan, Gary, 243, 252, 495, 539Lutzenberger, Hannah, 255 Luuk, Erkki, 258 Lyn, Heidi, 270 Mühlenbernd, Roland, 329 Müller, Thomas, 332 Maccagnani, Bettina, 519 Makuuchi, Michiru, 273 Manser, Marta, 101 Marelli, Marco, 83 Marghetis, Tyler, 513 Margiotoudi, Konstantina, 278 Martin, Jaycie, 525 Martin, Jose Segovia, 281 Martins, Mauricio, 80, 287 Martins, Pedro Tiago, 284 Matsuda, Tetsuya, 346 Matzinger, Theresa, 15 Mazuka, Reiko, 290, 501, 534 Meguerditchian, Adrien, 293 Meir, Irit, 295 Mengden, Ferdinand von, 232 Mengede, Janine, 516 Meyer, Antje, 402 Michaud, Jérôme, 306 Michelioudakis, Dimitris, 196 Micklos, Ashley, 316, 419 Mills, Gregory, 319 Moisik, Scott, 171 Moran, Steven, 322 Morin, Olivier, 332, 528 Morisseau, Tiffany, 332 Morrison, Robert, 150

Motamedi, Yasamin, 326 Murphy, Elliot, 335 Nölle, Jonas, 349 Nakamichi, Masayuki, 193 Niederhut, Dillon, 339 Nishimura, Takeshi, 346 Nomura, Yoshitaka, 346 O'Donnell, Timothy, 352 Ogura, Mieko, 355 Okanoya, Kazuo, 193, 364, 501, 504, 534 Oktaba, Katarzyna, 367 Onnis, Luca, 249 Papa, Aliki, 370 Parthemore, Joel, 372 Paul, Jing, 252 Pawlaczyk, Natalia, 50 Peña, Ines, 376 Peruffo, Sara, 519 Pickering, Martin, 458 Pijpops, Dirk, 379 Pleyer, Michael, 382 Poliva, Oren, 386 Prömer, Christina, 408 Pruetz, Jill D., 293 Pulvermüller, Friedemann, 278 Pyers, Jennie, 445 Quincoses, Charlotte, 445 Radkevich, Nina, 196 Ramsammy, Michael, 492 Ravignani, Andrea, 89, 399 Raviv, Limor, 402 Reinöhl, Uta, 95

Ridley, Amanda, 101

Ritt, Nikolaus, 15, 408

Rissman, Lilia, 405

Roberts, Gareth, 422, 475 Roberts, Sean, 67, 412, 419 Rolando, Sara, 425 Rossello, Joana, 428 Rudnicki, Konrad, 431 Sato, Asha, 492 Schlechtweg, Dominik, 434 Schouwstra, Marieke, 125, 212, 215, 326, 458, 492Seki, Yoshimasa, 437 Semenuks, Arturs, 439, 442 Semple, Stuart, 161 Senghas, Ann, 104, 445 Shillcock, Richard, 449 Shurley, Joleana, 458 Sievers, Christine, 461 Silva, Vinicius Macuch, 316 Silvey, Catriona, 464 Slocombe, Katie, 467 Smith, Andrew Marcus, 472 Smith, Kenny, 70, 111, 128, 158, 189, 215, 326, 352, 425,469, 531Sneller, Betsy, 475 Sonesson, Göran, 478 Spike, Matthew, 489 Stadje, Miriam, 370 Staib, Marlene, 349 Stevens, Jon, 422 Sulik, Justin, 246, 419, 495, 498 Suzuki, Reiji, 165 Takahasi, Miki, 501 Tamariz, Monica, 206, 281, 370 Thompson, Bill, 89 Tomoko, Mizuhara, 504 Toro, Juan M, 76 Torti, Valeria, 134 Townsend, Simon, 101 Toya, Genta, 507 Tylén, Kristian, 349

Ussishkin, Adam, 525 Valente, Daria, 134 Vallortigara, Giorgio, 519 Van de Velde, Freek, 379 Vasileva, Olga, 510 Verhoef, Tessa, 89, 513 Verkerk, Annemarie, 38, 322 Vernes, Sonja, 235, 238, 516 Versace, Elisabetta, 519 Villringer, Arno, 287 Vogt, Paul, 8 Vuillemin, Adelyne, 293 Wacewicz, Slawomir, 329, 548 Walde, Sabine Schulte Im, 434 Walker, Esther, 513 Wang, William, 355 Wedel, Andrew, 209 Wedel, Andy, 525 Wendorf, Ashley, 252 Whiten, Andrew, 5 Wiegrebe, Lutz, 235 Wilke, Claudia, 467 Willems, Erik P., 551 Winters, James, 332, 528 Woensdregt, Marieke, 531 Wonnacott, Elizabeth, 111 Yamada, Kazunori, 193 Yamane, Naoto, 534 Yasuda, Tetsuya, 230 Zürcher, Yvonne, 551 Zehentner, Eva, 537 Zeshan, Ulrike, 67 Zettersten, Martin, 539 Zlatev, Jordan, 548 Zola, Caroline, 445 Zuberbühler, Klaus, 461 Zywiczynski, Przemyslaw, 329, 548