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

STELLA FRANK[*1] and KENNY SMITH[1]

[*]Corresponding Author: stella.frank@ed.ac.uk
[1]Centre 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 non-native 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-

DOI:10.12775/3991-1.031

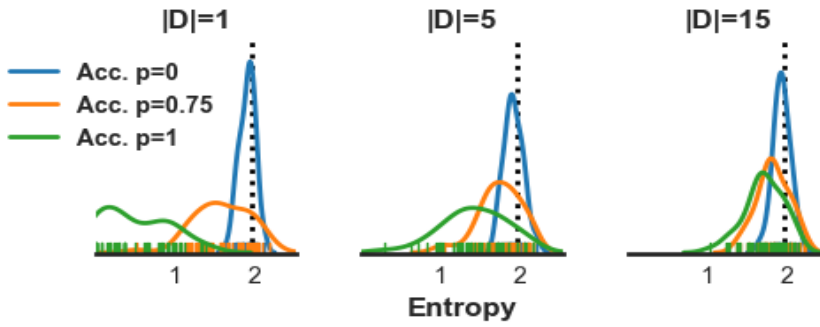

Figure 1. Distributions of final entropy of the non-native speaker's language after 100 interactions with a native speaker, for varying degrees of accommodation (indicated by colour) and different amounts of data encountered pre-interaction by the non-native speaker (indicated by the parameter |D|). These results are for a language where there are 5 variants available, whose probability distribution in the native speaker's language is (0.5,0.2,0.1,0.1,0.1); 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).

References

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.