

## DYNAMICAL PREFERENCES AND SELF-ACTUATED LANGUAGE CHANGES

JEROME MICHAUD<sup>\*1,2</sup>

<sup>\*</sup>Corresponding Author: jerome.michaud84@gmail.com

<sup>1</sup>Department of Sociology, Uppsala University, Uppsala, Sweden

<sup>2</sup>Department 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 of 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 S-shaped 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

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 if 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 of 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 agree 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 originate 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 (Real & 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 valued 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 *subjective* 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 it 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  $\pi$  of length  $V$  (the number of variants) and whose components take values be-

tween 0 and 1. This complements the frequency vector  $\mathbf{x}$  from which utterances  $\mathbf{u}$  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*  $\boldsymbol{\alpha} := \mathbf{1} - \boldsymbol{\pi}$  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

$$x_v^{(i),n+1} = x_v^{(i),n} + u_v^{(j),n} \sum_{w=1}^V \alpha_w^{(i),n} x_w^{(i),n} - \alpha_v^{(i),n} x_v^{(i),n}, \quad (1)$$

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  $\boldsymbol{\pi}$  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  $\mathbf{U}^{(i)}$  be the average utterance of the neighbors of an agent  $i$ . We define the updating rule for the preference vector  $\boldsymbol{\pi}^{(i)}$  by

$$\boldsymbol{\pi}^{(i),n+1} = \boldsymbol{\pi}^{(i),n} - \mu(\mathbf{x}^{(i),n} - \mathbf{U}^{(i),n}), \quad (2)$$

where  $\mu$  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  $\mathbf{U}^{(i),n} - \mathbf{x}^{(i),n}$ , 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  $\mathbf{x}^{(i)}$  with averaged utterance  $\mathbf{U}^i$  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  $\mu$  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.

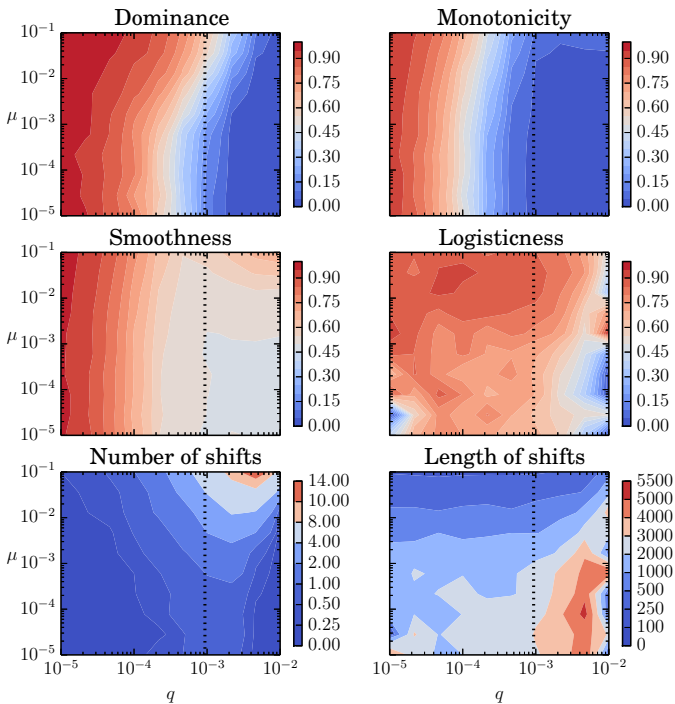


Figure 1. Results for the application of time series measures. For each measures, 10 logarithmically spaced values for  $q$  and  $\mu$  have been selected and, for each parameter combination, 100 runs of length 10000 network updates have been performed. The dashed vertical line corresponds to the critical  $q_*$  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  $\mu$ ). 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  $\mu$ . 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 *23rd 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.