Language Contact: Peaceful Coexistence or Emergence of a Contact Language

Jérôme Michaud
University of Edinburgh, SOPA
Peter Guthrie Tait Road
EH9 3FD Edinburgh, UK
jerome.michaud84@gmail.com

Gerhard Schaden
Université Lille SHS
CNRS UMR 8163 STL
59000 Lille
gerhard.schaden@univ-lille3.fr

Abstract

This paper presents a simple model of linguistic priming between languages in contact, based on the *utterance selection model* (USM) for language change of Baxter et al. (2006). It will be shown that the emergence or the non-emergence of a new contact language depends on the way potentially bilingual agents choose a language to communicate.

1 Introduction

One major factor driving language evolution is the interaction of its speakers. In our paper, we consider a situation where speakers of two different communities are in contact, and where (at least some of) the speakers of the two groups need to communicate with one another. There are basically two ways of resolving the communicative problem in such cases: speakers can either use (some variant of) their community languages or a contact language can emerge — which corresponds to neither of the two community languages. This new language, which can take the form of a pidgin, is not random and highly correlates with the two languages it originates from. The fine-grained processes controlling this process are poorly understood. In this paper, we provide a simple computational simulation of the stochastic dynamics of a contact situation. We show that the way agents choose a language when they interact partly controls the emergence of a contact language.

In order to capture the stochastic aspects of linguistic interactions, Baxter et al. (2006) designed the *utterance selection model* (USM) for language change (see also Croft, 2000). This is a stochastic agent-based model that accounts for the evolution of a single (socio-)linguistic variable (Tagliamonte, 2011), which can be instantiated in a finite number of equivalent variants. USMs can be seen as formal models of what Calvet (1999) calls an *ecolinguistic system*. The USM is well-adapted to capture the dynamics of a single linguistic variable and its stochastic evolution. It can also be used to predict the evolution of a linguistic variable in a larger population using coarse-graining techniques as shown in Michaud (2017). Other modelling methods, such as the model of Tria et al. (2015), can accurately reproduce the conditions under which a creole emerges based on census data. However, their model is highly idealized and makes assumptions such as "*if the hearer does not already possess the language of the utterance in her repertoire and therefore cannot make sense of it, she learns it by adding it to her repertoire*" (Tria et al., 2015, p. 6), which do not seem very realistic. Our aim is to provide a (still very simple) model, but whose agents correspond more closely to 'real' humans' capacities.

We study a simple extension of the USM that models potentially bilingual agents and explicitly takes into account a priming effect between the two languages to model a situation of language contact. In particular, we study how the choice of a specific language in the interaction can lead either to the coexistence of the two group languages or to the emergence of a new contact language.

2 Methodology

Our model is an extension of the USM for language change (Baxter et al., 2006) that takes into account potentially bilingual agents and models a priming effect between a group language and a non-group language. Below, we recall the definition of the USM and then explain the modifications made to model potentially bilingual agents. We conclude this section by explaining how we measure this stochastic system and explain how we decide when a contact language emerges.

2.1 The standard USM

The USM models the evolution of the usage frequency of a linguistic variable with V equivalent variants. The probability distribution over the different variants of an agent i is represented by the vector $\boldsymbol{x}^{(i)}$, where a component $x_v^{(i)}$ represents the probability/frequency with which agent i uses variant v.
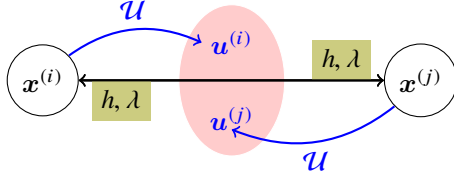

Figure 1: Structure of the USM interaction. On the chosen edge, the agents use their usage frequency vector $\boldsymbol{x}$ to produce an utterance $\boldsymbol{u}$ through the process $\mathcal{U}$, which depend on the matrix M. The utterances are then used to update the internal beliefs depending on the parameters h and λ (see below).

In order to communicate, an agent i produces an utterance $\boldsymbol{u}^{(i)}$ of length L from a production process $\mathcal{U}$ ($\boldsymbol{u} := \mathcal{U}\boldsymbol{x}$). The process $\mathcal{U}$ is defined by

$$\mathcal{U}\boldsymbol{x} = \frac{1}{L} M \mathrm{Multi}(L, \boldsymbol{x}), \qquad (1)$$

where M is a matrix representing production errors or innovations and $\mathrm{Multi}(L, \boldsymbol{x})$ is a vector counting the outcome of L multinomially sampled variables.

During an interaction, two connected agents are randomly selected. Then, they both produce an utterance $\boldsymbol{u}$ and update their usage frequency distribution $\boldsymbol{x}$ by

$$\boldsymbol{x}^{(i),\mathrm{new}} = \boldsymbol{x}^{(i),\mathrm{old}} + \delta\boldsymbol{x}^{(i)}, \qquad (2)$$

where the increment $\delta\boldsymbol{x}^{(i)}$ is defined by

$$\delta\boldsymbol{x}^{(i)} = \lambda[(1-h)\boldsymbol{u}^{(i)} + h\boldsymbol{u}^{(j)} - \boldsymbol{x}^{(i),\mathrm{old}}], \qquad (3)$$

where λ is a usually small learning parameter and the attention parameter h controls the relative importance of the utterance $\boldsymbol{u}^{(j)}$ of the other agents with respect to her own utterance $\boldsymbol{u}^{(i)}$. The presence of $\boldsymbol{u}^{(i)}$ accounts for a *self-monitoring* process and the presence of $\boldsymbol{u}^{(j)}$ accounts for an *accommodation* process. This learning rule assumes communicative success.[1]

---
[1]One way of interpreting this is to assume that the context and non-verbal communication provide sufficient clues for the interpretation.

The USM has been used to study the conditions under which a consensus can be achieved in a population (Baxter et al., 2006; Michaud, 2017). It has been applied to test the hypothesis of Trudgill about the emergence of New Zealand English (Baxter et al., 2009) and to test under which conditions the time series of usage frequency of an innovative variant takes the form of an S-shaped curve (Blythe and Croft, 2012).

2.2 Bilingual agents, social structure and priming

In order to model a language contact situation, the USM needs to take into account the possibility that agents become bilingual. We assume that each agent belongs to a group labelled by capital letters $(A, B, \dots)$ and every agent knows the group membership of every other agents. An agent belonging to some group Y is able to represent two languages and we denote the corresponding frequency vectors $\boldsymbol{x}_Y$ for the group language Y and $\boldsymbol{x}_{\bar{Y}}$ for the non-group language $\bar{Y}$. With this modification, the utterance production and learning rules have to be adapted.

During an interaction, if two agents belong to the same group, they interact as usual using the standard USM production and learning rules. If the two agents belong to different groups, we consider two scenarios:

Scenario 1: Symmetric adaptation When two agents of different groups interact, they both adapt to the other agent. For example if agent i of group A and agent j of group B interact, they both use their non-group language, i.e. $\bar{A}$ and $\bar{B}$, respectively.

Scenario 2: Unilateral adaptation When two agents of different groups interact, *for each interaction* they randomly choose a group language to use, either A (with probability p) or B (with probability $1 - p$), and the agent who doesn't know the language uses his non-group language. For example, if agent i belongs to group A and agent j belongs to group B, then one language is chosen randomly, say language of group A, then agent i uses her group language and j her non-group language $\bar{B}$.

When an agent uses her group language, her knowledge of the language is assumed to be perfect and she uses the corresponding frequency vector. However,

when an agent needs to use a non-group language, her knowledge is only partial and the non-group language is primed by the group language. This priming is implemented by the rule that whenever a non-group language has to be used, instead of using the frequency vector $\boldsymbol{x}_{\bar{A}}$ purely, the group language frequencies modifies the distribution through

$$\boldsymbol{x}_{\bar{A},\text{eff}} = (1 - \rho)\boldsymbol{x}_{\bar{A}} + \rho\boldsymbol{x}_A. \qquad (4)$$

The priming parameter ρ models the degree of mixing between languages A and $\bar{A}$. If $\rho = 0$, then there is no priming and the effective frequency vector boils down to $\boldsymbol{x}_{\bar{A}}$ and if $\rho = 1$, then priming is total and the effective frequency distribution $\boldsymbol{x}_{\bar{A},\text{eff}} = \boldsymbol{x}_A$. In the production rule (1), it is the effective frequency vector $\boldsymbol{x}_{\bar{A},\text{eff}}$ that is sampled. The learning rule (2) is the same but is only applied to the languages associated with the interaction.

The social structure used in our model is made of two random regular graphs of degree 3, containing 20 agents each, connected with each other by 5 connexions, see Fig. 2. The agents situated at an end of an intergroup connexion are the potentially bilingual agents, the other agents are monolingual, since they never use their non-group language.
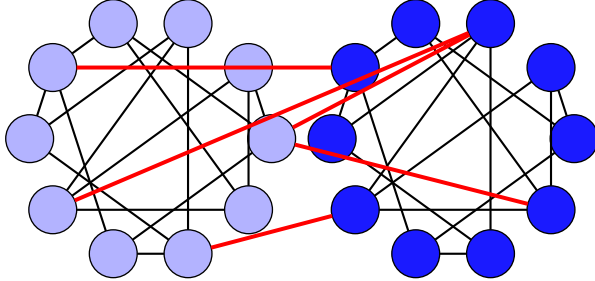

Figure 2: Illustration of the social structure. In the figure the groups have 10 agents instead of 20; red links are intergroup links.

2.3 Measuring the outcome of simulations

We measure the outcome of the simulation by computing Pearson's correlation coefficient between the time series of the averaged use of a language by each group. Note that the non-group languages are only used by agents with intergroup connexions and only these agents are updating their non-group language and can, therefore, become bilinguals.

We introduce the following notation: r_{XY} correlation between language X of group A and language Y of group B, illustrated in Fig. 3. If r_{XY} is close to 1, then the two languages can be considered
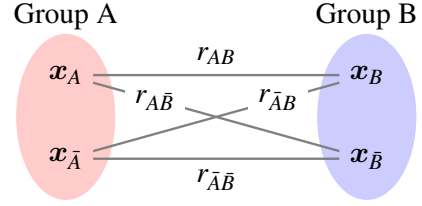

Figure 3: Illustration of the correlation coefficients between the different languages of the two groups.

as being the same. If r_{XY} is close to 0, the two languages are independent. For medium values of r_{XY} the languages are different but correlated.

3 Results

For the simulation of the two scenarios, we used the network topology discussed in Sec. 2.2 and illustrated in Fig. 2. The parameters are $N = 40$ agents with 5 intergroup connexions, the number of variants is $V = 3$, and the utterance length is $L = 2$. The learning parameter $\lambda = 0.1$ and the attention parameter $h = 0.5$. The matrix M used to simulate errors and innovations is of the form

$$M = \begin{bmatrix} 1-q & 0 & q \\ q & 1-q & 0 \\ 0 & q & 1-q \end{bmatrix}, \qquad (5)$$

where $q = 3 \times 10^{-4}$. The structure of this matrix is such that the innovations are ordered and variant 1 can only be transformed into variant 2, but not into variant 3, and similarly for the other variants. The pattern of mutation/innovation should be read columnwise. The simulations have been performed for $T = 5000$ full network updates and the priming parameter ρ is varied.

In Scenario 1, two interacting agents of different groups used their non-group language. Results are displayed in Fig. 4 and we observe that the correlation between $\boldsymbol{x}_A$ and $\boldsymbol{x}_B$ (r_{AB}) is close to zero for all values of the priming parameter ρ; the correlation between $\boldsymbol{x}_{\bar{A}}$ and $\boldsymbol{x}_{\bar{B}}$ ($r_{\bar{A}\bar{B}}$) is close to one for all values of the priming parameter ρ; the other correlation coefficients grow from 0 to about 0.7 when ρ is increased. From these results, one can conclude that there are three languages in these settings, the language of group A, the language of group B, and a new contact language $\bar{A} = \bar{B}$ partly correlated with both languages.

In Scenario 2, when two agents of different groups interact, at each interaction, they choose language A with probability p and language B
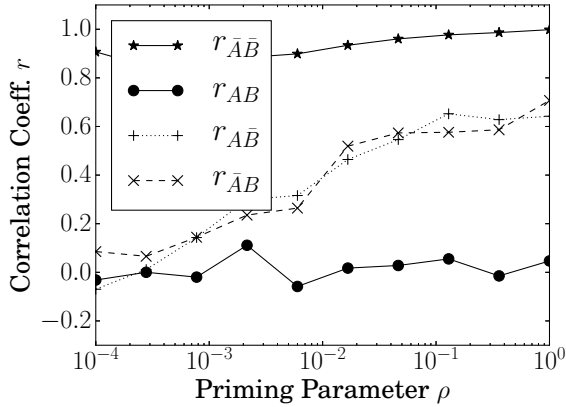
Figure 4: Scenario 1: Emergence of a contact language. The 3 groups of curves represent the different languages.
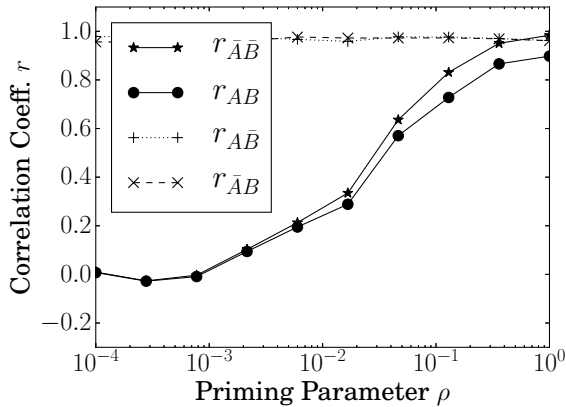

Figure 5: Scenario 2: No emergence of a contact language. The 2 groups of curves represent the different languages merging for high ρ.

with probability $1 - p$. Here $p = 0.5$ and the two languages are equivalent. Results are displayed in Fig. 5 and we observe that $r_{A\bar{B}}$ and $r_{\bar{A}B}$ are close to one for all values of ρ and the other correlation coefficients increase from zero to one as ρ increases. In this situation, there are only two languages present, the two group languages. When ρ is large enough, the two languages converge to the same language and there is a single language remaining.

4 Discussion

We have shown that the decision of which language to use has an important impact on the outcome of the simulation, and can lead either to the emergence of a contact language, or to the stable cohabitation of the two group languages. Compared to the naming

game model of Tria et al. (2015), the agents of our model do not instantaneously learn or forget a language but gradually adapt their behaviour. As a result, the emergence of a contact language, or absence thereof, is more gradual and better accounts for the influence of the stance that agents take during intergroup communication.

Our model makes a number of idealising assumptions. First of all, we assume that there is no reason to choose one language rather than the other for intergroup communication — which implies the absence of any hierarchy between the languages (or groups). This is probably a rather rare setting in the wild. There are different degrees of divergence from this configuration: instead of a perfectly symmetric situation with a probability $p = 0.5$ for using each language, there may be a different p tilted towards one group language. In extreme cases, if $p = 1$ or 0, or when the priming parameter $\rho = 1$, the agents of one group do not adapt to the language of the other group at all. Therefore, their group language will always be used, forcing the agents of the other group to adapt. Furthermore, in our model, the preferences and attitudes of the agents as well as the network structures do not evolve over time (bilinguals cannot switch group allegiance, etc.).

That being said, in which circumstances of real-life language contact would we expect the two scenarios we have considered to arise? Notice first that the asymmetric scenario should have a lower cognitive cost than the symmetric one, since only one agent in an intergroup interaction needs to adapt his behaviour, whereas scenario 1 requires both agents to do so. Using this argument, scenario 2 should be preferred overall and no contact language should emerge. One can also argue that an asymmetric scenario will take longer to reach a consensus through the population. As a consequence, if the pressure for communication is strong enough, the more costly, but more rapidly converging scenario 1 would be preferred and a contact language is likely to emerge. The additional cognitive cost of a symmetric adaptation should be partly compensated by the fact that contact languages are usually simpler than fully-fledged languages.

To conclude, scenario 1 is expected if communication pressure is strong and the group languages are unrelated. Otherwise, we expect scenario 2. This is consistent with the conclusions of Tria et al. (2015) concerning the influence of population structure on communicative needs and creole-formation.

Acknowledgments

References

Gareth J. Baxter, Richard A. Blythe, William Croft, and Alan J. McKane. 2006. Utterance selection model of language change. *Physical Review E*, 73(4):046118.

Gareth J. Baxter, Richard A. Blythe, William Croft, and Alan J. McKane. 2009. Modeling language change: An evaluation of Trudgill's theory of the emergence of New Zealand English. *Language Variation and Change*, 21(02):257–296.

Richard A. Blythe and William Croft. 2012. S-curves and the mechanisms of propagation in language change. *Language*, 88(2):269–304, June.

Louis-Jean Calvet. 1999. *Pour une écologie des langues du monde*. Plon, Paris.

William Croft. 2000. *Explaining language change: An evolutionary approach*. Pearson Education.

Jérôme Michaud. 2017. Continuous time limits of the utterance selection model. *Phys. Rev. E*, 95:022308, Feb.

Sali A. Tagliamonte. 2011. *Variationist sociolinguistics: Change, observation, interpretation*, volume 39. John Wiley & Sons.

Francesca Tria, Vito D.P. Servedio, Salikoko S. Mufwene, and Vittorio Loreto. 2015. Modeling the emergence of contact languages. *PloS one*, 10(4):e0120771.