<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper presented at *COMPLEX NETWORKS 2019, Lisbon, December 10-12, 2019.*

Citation for the original published paper:

Michaud, J. (2020)

Modelling Opinion Dynamics and Language Change: Two Faces of the Same Coin
In: *Complex Networks and Their Applications VIII: Volume 2 Proceedings of the Eighth International Conference on Complex Networks and Their Applications COMPLEX NETWORKS 2019* (pp. 305-315). Springer
https://doi.org/10.1007/978-3-030-36683-4_25

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-398008>

Modelling opinion dynamics and language change: two faces of the same coin

Jérôme Michaud¹

Department of Sociology, Uppsala University,
Thunbergsvägen 3H, 752 38 Uppsala, Sweden,
`jerome.michaud@soc.uu.se`

Abstract. While opinion dynamics models and language change models seem to be very different at first sight, we demonstrate in this paper that they are more similar than they look like. Here, we analyse the similarities and differences between the Social Influence with Recurrent Mobility model (Phys. Rev. Lett. **112**, 158701) and the Utterance Selection Model of language change (Phys. Rev. E **73**, 046118), two models that simulate a stochastic dynamics on a complex network, and show that their mesoscopic dynamic is strikingly similar. By drawing an analogy between the two models, we discuss possibilities of cross-fertilization of research between the two problems.

Keywords: Social Influence with Recurrent Mobility, Utterance Selection Model, Stochastic dynamics on network, mesoscopic approximation, Stochastic differential equation, multiplicative noise

1 Introduction

At a first sight, opinion dynamics and language change seem to be very different topics and similarities between them are not evident. This paper aims at exemplifying such similarities between these two problems by looking at stochastic models of opinion dynamics and language change on complex networks. The analysis justifies to make an analogy between the two problems and enables a reinterpretation of the dynamics of both problems. As a result, the models for opinion dynamics and language change can be seen at two faces of the same modelling coin and this duality can be utilized to improve both models.

The chosen model for opinion dynamics is the Social Influence with Recurrent Mobility model [5, 10], which is a generalization of the well-studied Voter Model [7]¹ and the chosen model for language evolution is the Utterance Selection Model (USM) for language change [1], which is a model grounded in evolutionary theory and inspired by the Wright-Fischer model for population genetics [12, 6]. These two models will be shown to have a very similar dynamics, at least at some mesoscopic scale.

¹ But note that the SIRM model relaxes some of the problematic assumptions of the Voter Model. Errors can occur in the copying processes and unilateral change of opinion is possible.

This paper is organized as follows. Section 2 introduces the two studied models and their respective mesoscopic approximation that takes the form of a stochastic differential equation (SDE) with multiplicative noise. Section 3 discusses the similarities and differences between the two models. Section 4 provides some concluding remarks and suggests directions in which the models studied could be improved.

2 Two stochastic models of opinion dynamics and language change on networks

The two models we are analyzing in this paper are stochastic models on networks whose nodes can be partitioned into a finite number of groups. The level at which the models will be compared is at this group level, where the mesoscopic dynamics take a very similar form.

The SIRM and the USM share the property of having two sources of stochasticity: the standard stochasticity originating from random ordering of pairwise interactions and another level of stochasticity originating in imperfect interactions. One specificity of this second source of stochasticity is that it survives a (heterogeneous-) mean-field treatment, leading to a stochastic dynamics of the mesoscale quantities.

2.1 The Social Influence with Recurrent Mobility (SIRM) model

The SIRM model [5] is a generalization of the well-studied Voter Model [7] that considers the evolution of the opinions of a population partitioned into geographic/administrative regions. The coupling between the regions is encoded using the commuting pattern of the individuals, i.e., using the number of people living in one region and working in another. It has been applied to US presidential elections [5] and to Swedish parliamentary elections [10]. The formulation we are using is that of [10], where some issues in the original mathematical formulation of the SIRM have been fixed.

Network structure Consider a population of N individuals divided into M regions usually representing electoral regions, such as counties, municipalities or states. Let N_{ij} be the number of commuters between regions i and j , that is the number of individuals living in region i and working in region j . The commuting network is constructed by taking the M municipalities as nodes and linking them by weighted edges of weight N_{ij} . The commuting network is a directed weighted network with self-loops, since there are people who live and work in the same place. In the rest of the paper, we will refer to the sub-population N_{ij} as the *commuting cell* ij .

Additionally, taking the commuting cells as nodes, one can define the *recurrent mobility network* in which each node has an attribute N_{ij} and is connected to all others commuting cells sharing either the first or the second index with

them. That means that all commuting cells $i\cdot$ or $\cdot j$ are connected to the commuting cell ij . The recurrent mobility network can be partitioned into regions by grouping the commuting cells according to their first index.

From the N_{ij} quantities, one can construct the number N_i of people living in region i by summing over the second index and the number N'_j of people working in region j by summing over the first index. Hence, we define

$$N_i := \sum_j N_{ij} \quad \text{and} \quad N'_j := \sum_i N_{ij}. \quad (1)$$

The total population $N = \sum_i N_i = \sum_j N'_j$. The primed quantities refer to the working population.

Opinion structure and vote shares Let us assume that individuals can choose between K different opinions. We denote by V_{ij}^k the number of people in commuting cell ij that have opinion k . For consistency, we must have

$$\sum_{k=1}^K V_{ij}^k = N_{ij}. \quad (2)$$

The state of the model is fully defined by the quantities V_{ij}^k . For convenience, we also introduce the vote shares

$$v_{ij}^k := \frac{V_{ij}^k}{N_{ij}}. \quad (3)$$

With this definition one can easily show that

$$v_i^k = \sum_j \frac{N_{ij}}{N_i} v_{ij}^k \quad \text{and} \quad v'_j = \sum_l \frac{N_{lj}}{N'_j} v_{lj}^k \quad (4)$$

are the vote shares of opinion k for the population living in region i and for the population working in region j .

Dynamics Let us now define the transition operators $R_{ij}^{kk'}$

$$R_{ij}^{kk'} := P[(V_{ij}^k, V_{ij}^{k'}) \rightarrow (V_{ij}^k - 1, V_{ij}^{k'} + 1)] = \frac{N_{ij}}{N} v_{ij}^k p_{ij}^{k \rightarrow k'}, \quad (5)$$

that defines the probability that an individual changes from opinion k to opinion k' in the commuting cell ij . For the left most term, the first factor is the probability to choose an individual in the commuting cell ij ; the second factor is the probability that this individual holds opinion k and the third factor is the probability that she changes opinion to k' .

A procedure to get a stochastic version of a probability distribution using a Dirichlet distribution has been developed in [10]. This procedure can be utilized

to add stochasticity to the rates of the SIRM model. In the rest of this paper, we consider the following model:

$$p_{ij}^{k \rightarrow k'} = \lambda \left(\alpha v_i^{k'} + (1 - \alpha) v_j'^{k'} \right) + \beta \tilde{v}_{ij,D}^{k'} + \gamma / K, \quad (6)$$

where $\lambda, \beta, \gamma \geq 0$ and satisfy $\lambda + \beta + \gamma = 1$ and where

$$\tilde{\mathbf{v}}_{ij,D} = \text{Dir}(\mathbf{v}_{ij}/D), \quad (7)$$

where $\text{Dir}(\mathbf{v}_{ij}/D)$ denotes a sample from the Dirichlet distribution of parameter $\mathbf{v}_{ij}/D$ and D controls the level of noise associated with the intra-commuting-cell influence. The parameter β represents the strength of the intra-commuting-cell influence and mainly act as a source of stochasticity, the parameter γ controls the intensity of free will encoding unilateral change of opinion, where all opinions are equally likely to be chosen. One parameter can be eliminated by rescaling the time. The state of the model is updated by resampling the population of each commuting cells according to the probabilities (6) as described in [10].

Mesoscale approximation A continuous time limit can be derived for the evolution of the commuting cell vote shares v_{ij} . As obtained in [10], it takes the form

$$dv_{ij} = \hat{d}(v_{ij})dt + \sqrt{\hat{D}(v_{ij})}dW_{ij}^*(t), \quad (8)$$

where $dW_{ij}(t)$ is a white noise, $\hat{d}(v_{ij})$ is a drift function and $\hat{D}(v_{ij})$ are diffusion coefficient.

In the case of two variants with the transition probability defined in (6), the drift and diffusion coefficients, $\hat{d}(v_{ij})$ and $\hat{D}(v_{ij})$, take the form

$$\begin{aligned} \hat{d}(v_{ij}) &= \lambda[\alpha v_i + (1 - \alpha)v_j' - v_{ij}] + \beta(\tilde{v}_{ij,D} - v_{ij}) + \gamma \left(\frac{1}{2} - v_{ij} \right) \\ \hat{D}(v_{ij}) &= \frac{1}{N_{ij}} \left[(1 - 2v_{ij}) [\lambda[\alpha v_i + (1 - \alpha)v_j'] + \beta \tilde{v}_{ij,D}] + (1 - \gamma)v_{ij} + \frac{\gamma}{2} \right]. \end{aligned} \quad (9)$$

These two functions are stochastic functions because of the presence of $\tilde{v}_{ij,D}$ in the formulation, which is a stochastic variable coming from the Dirichlet sampling process (7). As in previous analysis of the SIRM model, we neglect the diffusion function $\hat{D}(v_{ij})$ and use the normal approximation of Dirichlet distributed variable to obtained the SDE for the evolution of v_{ij}

$$\begin{aligned} dv_{ij} &= \left\{ \lambda [\alpha v_i + (1 - \alpha)v_j' - v_{ij}] + \gamma \left(\frac{1}{2} - v_{ij} \right) \right\} dt \\ &\quad + \beta \sqrt{\frac{\tilde{D}}{\tilde{D} + 1}} \sqrt{v_{ij}(1 - v_{ij})} dW_{ij}. \end{aligned} \quad (10)$$

The mesoscopic approximation of the dynamics is then given by the evolution of v_i , which can be obtained by combining (4) with (10). Reorganizing the terms,

we obtain the following SDE

$$\begin{aligned} dv_i = & \left[C_1^{\text{SIRM}} \left(\frac{1}{2} - v_i \right) + C_2^{\text{SIRM}} \left(\sum_j \frac{N_{ij}}{N_i} v'_j - v_i \right) \right] dt \\ & + C_3^{\text{SIRM}} \sum_j \frac{N_{ij}}{N_i} \sqrt{v_{ij}(1 - v_{ij})} dW_{ij}, \end{aligned} \quad (11)$$

where the coefficients are given by

$$C_1^{\text{SIRM}} = \gamma, \quad (12)$$

$$C_2^{\text{SIRM}} = \lambda(1 - \alpha), \quad (13)$$

$$C_3^{\text{SIRM}} = \beta \sqrt{\frac{D}{D+1}}. \quad (14)$$

Equation (11) will be used as a basis for comparing the SIRM model with the USM.

2.2 The Utterance Selection Model (USM) for language change

The USM for language change [1] is an evolutionary model of language change inspired by population genetics. It has been shown in [4] that it is formally equivalent to the Wright-Fischer model of population genetics, where the population is divided into islands and where migration is possible between the islands. In this analogy, every speaker is an island and variants compete for being used. Variants are transmitted from one speaker to another through conversation and the production of utterances, which corresponds to the migration process between islands. This model and its extensions have been used to test hypothesis on new language formation [2], to explain sociolinguistic patterns [3] or to provide a theory for the self-actuation of changes in the language dynamics [11, 8].

Network structure The USM describes the evolution of N speakers represented as a node of a network. The network of speakers is static and a parameter h_{ij} is associated with the edge connecting speaker i to speaker j . This parameter controls the attention that speaker i pays to speaker j , which is not necessarily symmetrical. For simplicity, this parameter will be assumed not to depend on the identity of the speakers $h_{ij} = h$ for all i, j . This parameter plays a similar role than the α parameter of the SIRM model as it controls the importance of neighbors on the network.

We will further assume that speakers are partitioned into M communities. These communities can be defined as any group of speakers, but the heterogeneous mean-field approximation developed in [9] and used in this paper works better if the nodes associated with the agents are well-connected. For each community c , the number of speakers in that community is denoted by N_c and the averaged node degree of that community is denoted by k_c .

State vector Every agent i is characterized by a state vector $\mathbf{x}_i$ of length V representing the discrete probability distribution of using one of the available V variants, i.e., x_i^v is the probability that agent i uses variant v .

For each community c , the average state vector $\mathbf{x}_c$ is defined by

$$\mathbf{x}_c := \frac{1}{N_c} \sum_{j \in c} \mathbf{x}_j, \quad (15)$$

where N_c is the number of agents in community c and the sum runs over the member of the community.

Dynamics The dynamics of the USM is driven by pairwise interactions along the edges of the network. At each step a speaker i and a speaker j exchange utterances $\mathbf{u}$ defined as a biased sample of their state vector $\mathbf{x}$, i.e.,

$$\mathbf{u} := \frac{1}{L} M \text{Multi}(L, \mathbf{x}), \quad (16)$$

where L is the length of the utterance, M is an innovation matrix (row stochastic) and $\text{Multi}(L, \mathbf{x})$ denotes a multinomial sample of length L and parameter $\mathbf{x}$, and update their state $\mathbf{x}$.

The change in state vectors of the USM [1, 9] is given by

$$\delta \mathbf{x}_i = \lambda [(1 - h)(\mathbf{u}_i - \mathbf{x}_i) + h(\mathbf{u}_j - \mathbf{x}_i)], \quad (17)$$

where λ is a learning parameter and h is the attention parameter controlling the weight of the incoming utterance from speaker j with respect to speaker i own utterance. Equation (17) fully determines the dynamics of the USM.

Mesoscopic approximation A *stochastic heterogeneous mean-field* (SHMF) approximation of the USM has been developed and analysed in [9]. In order to compare the dynamics of the SIRM model with that of the USM, we provide the SHMF of the USM for two variants. In this case, state vectors take the form

$$\mathbf{x} = \begin{pmatrix} x \\ 1 - x \end{pmatrix}, \quad (18)$$

and the equation for the first component is sufficient to characterize the dynamics.

In the utterance production rule (16), an innovation matrix M should be specified. For this analysis, we use the symmetrical case and set

$$M := \begin{bmatrix} 1 - q & q \\ q & 1 - q \end{bmatrix}, \quad (19)$$

where q is an innovation parameter.

With these specifications, the SHMF approximation for the quantities x_c describing the communities dynamics is given by

$$\begin{aligned} dx_c = & \left[C_1^{\text{USM}} \left(\frac{1}{2} - x_c \right) + C_2^{\text{USM}} \left(\sum_{c'} p(c'|c) x'_{c'} - x_c \right) \right] dt \\ & + C_3^{\text{USM}} \left[(1-h) \sqrt{\frac{x_c(1-x_c)}{k_c N_c}} dW_c + h \sum_{c'} p(c'|c) \sqrt{\frac{x_{c'}(1-x_{c'})}{k_{c'} N_{c'}}} dW_{c'} \right], \end{aligned} \quad (20)$$

where dW_c and $dW_{c'}$ are white noises, and $p(c'|c)$ is the probability that a speaker in community c interacts with a speaker in community c' , which can be computed by computing the fraction of edges originating in community c that terminate in community c' . For consistency, we have

$$\sum_{c'} p(c'|c) = 1. \quad (21)$$

The variable $x'_{c'} := (M\mathbf{x}_{c'})_1$ is the first component of the biased state vector $M\mathbf{x}_{c'}$ and is given by

$$x'_{c'} = x_{c'} - 2q \left(\frac{1}{2} - x_{c'} \right). \quad (22)$$

The coefficients of (20) are given by

$$C_1^{\text{USM}} = \lambda k_c 2q(1-h), \quad (23)$$

$$C_2^{\text{USM}} = \lambda k_c h, \quad (24)$$

$$C_3^{\text{USM}} = \lambda k_c (1-2q) \frac{1}{\sqrt{L}}, \quad (25)$$

where λ is a learning parameter, k_c is the averaged node degree of community c , h the attention parameter, q the innovation parameter and L is the length of utterances. Equations (11) and (20) will be compared and discussed in the next section.

3 Analogy between the SIRM and the USM

In the previous section, we have derived mesoscopic approximations of both the SIRM model and the USM. These two equations (11) and (20) are strikingly similar, but also have some significant differences that we will now explore. We start by comparing the deterministic parts (first line of (11) and (20)) and then discuss the stochastic parts.

3.1 Deterministic terms

The deterministic parts of (11) and (20) contain two terms: a term due to free will/innovation and a term encoding the effect of interactions. We now discuss these two components.

The free will/innovation term takes the form $C_1(\frac{1}{2} - x)$ in both models and only differs by the form of the C_1 constant. In the SIRM model, this constant is a model parameter, whereas in the USM it takes a complex form given by (23). Assuming that the γ parameter of the SIRM model is usually small, and that the innovation parameter q of the USM is also a small quantity, we can argue that these two parameters play the same role; that of controlling the strength of unilateral change of opinion or variant either by free will or by innovation. As a result, we say that γ corresponds to q and encodes *unilateral change*.

The interaction term of (11) and (20) encodes the average influence of the other groups. Both terms are of the form $C_2 \left(\sum_j w_j x'_j - x_i \right)$, where the sum is the weighted average of the importance of neighboring groups.

The weights w_j encodes in both case topological features of the underlying networks. In the SIRM, $w_j := \frac{N_{ij}}{N_i}$ weight the importance of a neighboring region j by the fraction of people living in i commuting there. In the USM, $w_{c'} := p(c'|c)$ computes the fraction of people in community c talking to people in community c' . Therefore, we say that $\frac{N_{ij}}{N_i}$ corresponds to $p(c'|c)$ and encodes the *relative weight of neighbors*.

In the SIRM, the incoming signal from neighbors is encoded by $x'_j := v'_j$ the vote share at work in region j , whereas in the USM it is encoded by $x'_{c'}$, which represents the perceived utterance from community c' . The forms of these two terms are different and depend on the definition of the models. In the SIRM, the vote share at work in region j is given by (4) and introduce a coupling with all regions l for which $N_{lj} \neq 0$. Such a coupling between multiple regions is absent from the USM in which the incoming signal is simply a biased version of the state vector of the community and is given by (22). Such a difference can be interpreted by the fact that communizing cells encode specific behavior between regions i and j . In the language change context of the USM, this would correspond to adapting the utterances to the interlocutor. This feature is absent from the USM and could be added to it using “commuting cell” corresponding to state vectors tailored to the interlocutor. In addition, the biased aspect of the incoming signal of the USM encoded in the matrix M is not encoded in the incoming signal, but as an external noise. This difference explains the absence of the γ parameter in the incoming signal of the SIRM model. For the analogy, we say that v'_j corresponds to $x'_{c'}$ and encodes the *neighbor signal*.

Looking at the C_2 constant (13) and (24), we see that the parameter $1 - \alpha$ in the SIRM plays a similar role to the parameter h in the USM. Both weight the importance of interactions with respect to other processes. Therefore, we say that $1 - \alpha$ in the SIRM corresponds to the attention parameter h in the USM representing the *importance of interactions*.

3.2 Stochastic terms

The stochastic terms in (11) and (20) take a slightly different form, that originates in where uncertainty in interactions is introduced. In the SIRM model,

incertainty is added into the intra-commuting cell term $\beta \tilde{v}_{ij,D}^{k'}$ in (6). This formulation enables to recover the original formulation of the SIRM as a limit [5]. In the USM, however, the incertainty in interactions comes from the definition of utterances (16) and affect both utterances in the update rule (17). This difference is at the origin of the different form of the stochastic terms. However, the two formulations share the property to have multiplicative noise terms of the form $\sqrt{x(1-x)}dW$ originating either from a multinomial or a Dirichlet (Beta for two opinions) sampling process. In the SIRM, these multiplicative noises are indexed to the commuting cell ij , whereas in the USM they are indexed to the communities and not to subpart of them. It is possible to modify the dynamics of the SIRM to ressemble even more that of the USM by changing (6) to

$$p_{ij}^{k \rightarrow k'} = \lambda \left(\alpha \tilde{v}_{i,D}^{k'} + (1 - \alpha) \tilde{v}_{j,D}^{k'} \right) + \gamma/K, \quad (26)$$

where uncertainty is added to the vote shares at home and at work. Furthermore, since the uncertainty is included at the group level in the USM, the prefactors of noise are also dependent on the connectivity and size of the groups, whereas in the SIRM, the noise is added at the level of commuting cells, which are taken as nodes of the recurrent mobility network. This explains why in (11) there is no dependence of topological quantities in the noise term.

Despite these differences in structure, there is a similarity in the constant C_3 that enables to further extend the analogy between the two models. For instance, the parameter D in the SIRM model and the parameter L in the USM both controls the strength of the uncertainty in the incoming signal through a sampling process. It turns out that the covariance matrix of multinomial and Dirichlet processes only differ by a constant, since the Dirichlet distribution is the continuous analog to the multinomial distribution. Therefore, there is a correspondance between $\frac{D}{D+1}$ in the SIRM model and $\frac{1}{L}$ in the USM and encodes a *noise parameter*.

4 Discussion and outlook

As we have shown, the SIRM model and the USM for language change are deeply related. The correspondance between the SIRM model and USM is summarized in Table 1.

This correspondance stresses the similarities between the two models and illustrates the fact that these two approaches are two faces of the same modelling coin, that of modelling stochastic dynamics on complex networks. Since their formulations comes from different lines of research they also have some significant differences. For example, the working population of a region i is made of contributions from many different regions whose vote shares feed back to region i . This long range coupling is absent from the USM, in which every speaker contributes either to the region where they live or where they work.

The differences between the two models suggest ways to develop and improve both models. For example, the USM for language change has the property that

Table 1. Correspondance between the parameters of the SIRM model and of the USM.

	SIRM	USM
Unilateral change	γ	q
Neighbor signal	v'_j	$x'_{c'}$
Relative weight of neighbors	$\frac{N_{ij}}{N_i}$	$p(c' c)$
Importance of interactions	$1 - \alpha$	h
Noise parameter	$\frac{\bar{D}}{1+\bar{D}}$	$\frac{1}{L}$

speakers use the same state vector $\mathbf{x}$ in all their interactions, whereas the coupling between region i and region j is specific to this particular pair of regions and encoded through $\mathbf{v}_{ij}$. Such a substructure could be added to the USM to encode adaptation of speech to the identity of the neighbors and make this model more realistic. Furthermore, the SIRM model could be modified to incorporate more naturally the different sources of noise and, therefore, have a more elegant form. This could be done by adding noise on the vote shares at home and at work as in (26) and by assuming that innovation or errors can occur at this level by, for example, redefining the quantity $\tilde{\mathbf{v}}_{ij,D}$ to be

$$\tilde{\mathbf{v}}_{ij,D} = M\text{Dir}(\mathbf{v}_{ij}/D), \quad (27)$$

where M is a mutation matrix similar to that of the USM.

In conclusion, we have shown in this paper how the SIRM model and the USM are similar in many ways and how their differences can lead to improvements of both models. Since the USM has connections to evolutionary models, one can further the analogy presented here to interpret opinion dynamics models such as the Voter Model or the SIRM model in evolutionary terms. Furthermore, extensions of the USM have been developed [3, 11, 8] and the analogy developed here can be extended as well, opening new avenues of research.

References

1. Baxter, G.J., Blythe, R.A., Croft, W., McKane, A.J.: Utterance selection model of language change. *Physical Review E* 73(4), 046118 (2006), <https://doi.org/10.1103/PhysRevE.73.046118>
2. Baxter, G.J., Blythe, R.A., Croft, W., McKane, A.J.: Modeling language change: An evaluation of Trudgill’s theory of the emergence of New Zealand English. *Language Variation and Change* 21(02), 257–296 (2009), <https://doi.org/10.1017/S095439450999010X>
3. Baxter, G.J., Croft, W.: Modeling language change across the lifespan: Individual trajectories in community change. *Language Variation and Change* 28(02), 129–173 (2016), <https://doi.org/10.1017/S0954394516000077>
4. Blythe, R.A., McKane, A.J.: Stochastic models of evolution in genetics, ecology and linguistics. *Journal of Statistical Mechanics: Theory and Experiment* 2007(07), P07018 (2007), <https://doi.org/10.1088/1742-5468/2007/07/P07018>

5. Fernández-Gracia, J., Suchecki, K., Ramasco, J.J., San Miguel, M., Eguíluz, V.M.: Is the voter model a model for voters? *Phys. Rev. Lett.* 112, 158701 (2014), <https://link.aps.org/doi/10.1103/PhysRevLett.112.158701>
6. Fisher, R.A.: *The genetical theory of natural selection*. Clarendon Press, Oxford (1930)
7. Holley, R.A., Liggett, T.M.: Ergodic theorems for weakly interacting infinite systems and the voter model. *Ann. Probab.* 3(4), 643–663 (1975), <https://doi.org/10.1214/aop/1176996306>
8. Michaud, J.: Dynamic preferences and self-actuation of changes in language dynamics. *Language Dynamics and Change* 1, 61–103 (2019), <https://doi.org/10.1163/22105832-00901003>
9. Michaud, J.: Continuous time limits of the utterance selection model. *Phys. Rev. E* 95, 022308 (2017), <https://link.aps.org/doi/10.1103/PhysRevE.95.022308>
10. Michaud, J., Szilva, A.: Social influence with recurrent mobility and multiple options. *Phys. Rev. E* 97, 062313 (2018), <https://link.aps.org/doi/10.1103/PhysRevE.97.062313>
11. Stadler, K., Blythe, R.A., Smith, K., Kirby, S.: Momentum in language change. *Language Dynamics and Change* 6(2), 171–198 (2016), <https://doi.org/10.1163/22105832-00602005>
12. Wright, S.: Evolution in mendelian populations. *Genetics* 16(2), 97 (1931)