A model for phonetic changes driven by social interactions

A. Chacoma, N. Almeira, J.I. Perotti, O.V. Billoni

Abstract

We propose a model to study phonetic changes as an evolutionary process driven by social interactions between two groups of individuals with different phonological systems. Particularly, we focus on the changes in the place of articulations, inspired by the drift /f/ → /h/ observed in some words of Latin root in the Castilian language. In the model, each agent is characterized by a variable of three states \( S = \{1, 2, 3\} \), representing the place of articulation used during speech production. In this frame, we propose rules of interactions among agents which lead to phonetic imitation and consequently to changes in \( S \). Based on this, we mathematically formalize the model as a problem of population dynamics, derive the equations of evolution in the mean field approximation, and study the emergence of three non trivial global states, which can be linked to the pattern of phonetics changes observed in the language of Castile and in other Romance languages.

1. Introduction

Oral communication, as the process of transmitting concepts and ideas from one individual to another by word of mouth, has been a main feature of human kind since the first primitive societies. Historically, research in this field has been faced by anthropologists and linguistics, however, in the last years the interest for the development of new technologies related to automatic speech recognition, specially in artificial intelligent systems [3], has made it an active multidisciplinary area of research [36, 14, 13, 32].

Speech production might be thought as the combination of several cognitive processes: the selection of the proper words to express an idea, the suitable choice of a grammatical form, and the production of sounds via the motor system and the vocal apparatus [22]. In this work we focus on the latter, therefore, in the following we describe the most relevant concepts regarding the production of sounds.

Formally, phonemes are the minimal units of either vocalic or consonant sounds needed to produce words. In this regards, the set of phonemes which encompasses all the sounds needed to produce every word in a given language,
define a phonological system (PS). It is particularly important to emphasize, phonemes are not sounds but formal abstraction of speech sounds. Any phoneme in a PS might be a representation for a family of sounds, technically called phones which are recognized by speakers and linked to a specific sound during oral communications [30]. In this sense, the set of phones linked to a particular phoneme is known as allophone. For instance in English, the phoneme /t/ can be listened normal like in take /teIk/, similar to d in water /waːtər/, like a glottal stop in kitten /kɪtɪn/, or become silent like in winter /ˈwɪntər/. Physiologically, the process by which the vocal apparatus produces sounds is called phonation. This process naturally occurs during the breath when air is expelled from the lungs through the glottis, creating a pressure drop across the larynx. Under this conditions, every phone may be produced by three basic mechanisms, (i) vocal folds vibration, (ii) total or partial air flow interruptions, and (iii) a combination of i and ii. By phonation, humans are able to produce a wide range of sounds, usually divided into two groups: Vowels and consonants [20]. Let us focus on consonants production. In this case, the phonetic apparatus uses a combination of tongue, lips, teeth and the soft palate, in order to shape the different air obstructions needed to produce the sounds. The point inside the vocal cavity where the obstruction occurs is called articulation place (AP), and the manner in which the obstruction is shaped is called articulation mode (AM). This two dimensions, AP and AM, are commonly used to classify consonants, as well as we show in table 1 for the case of Castilian language. Symbols in the table are graphic representations of phonemes, expressed by the standardized international phonetic alphabet (IPA) [34].

On the other hand, since languages are in continuous evolution, it is well known that under certain conditions, this process might lead to changes into the PS [18]. Particularly, in this work we are interested in studying phonetic changes in consonants related to variations in the AP. In this regard, it has been observed that these changes are enhanced when two or more groups of people with different languages are forced to socially interact [21, 17, 7, 5, 4]. For instance, when a group invades others group’s territory, or when two groups establish economic relations (trade, exchange of services, etc). The linguistic, phonetic and grammatical mutual influences produced by the interaction among the groups, define a linguistic stratum (LS) [19, 35], where the persistent social interaction over time, in a process of oral communication, guides the evolution to a common PS and to a common new language.

A notable example of phonetic change in the AP due to a LS, and the main inspiration of this work, is the case of the glottalization of the bilabial fricative phoneme /φ/ towards the glottal fricative /h/ (hereafter referred to as the change /φ/→/h/), and its subsequent disonorisation in some Latin root words of Castilian language (see table 2). It is thought that, the social process which led to the LS in this particular case, was related to the social interactions among the prehispanic tribes (Iberics, Asturians, and mainly Vascons) and the Romans, which were forced to socially interact in the Iberian peninsula— during the period of Rome’s domain— from the second to the ninth century, AD [23, 1, 6]. In this particular LS there were groups of people with a PS (based on Celtic language), socially interacting with other group of people with a total different PS (based on Latin language).

In this context, it is thought the change /φ/→/h/ is related to prehispanic tribes speakers performing changes in the AP during fricatives productions,
trying to improve their communication skills with Latin speakers [15]. Notably, these changes are not observed in other Romance languages on the Iberian peninsula, as the case of Portuguese or Catalan, which emerge from a similar LS than Castilian. This fact led to researchers to theorize about the properties of this particular LS, and additionally to propose alternative theories [16, 29]. Until now, it seems there is not a total consensus regarding the causes which led to these different evolution of the Romance languages, therefore the case is currently considered by linguistics as an open problem.

In the context of what we have said above, the main idea of this work is modeling a LS based on two groups with different PS which change the AP guided by social interaction rules based on phonetic imitation. We propose to face the problem in the frame of population dynamics, where we study the evolution of the changes in both groups, and the emergence of general states of pronunciation in the LS. The paper is divided into three main sections: in section 2 we mathematically formalize the model, define the main variables, propose the rules of the interactions and describe the dynamics; in section 3 we derive from first principles the equations of evolution; and, lastly, in section 4 we analyze the emergence of global states by performing a stability analysis of the evolution equations. We found that by tuning the parameters related to social interactions in the LS, our model shows the emergence of three general global states which capture well the observations of the emergent Romance languages of the Iberian peninsula, reinforcing from our mathematical approach, the stratum-based theories present in the literature.

2. The model

We aim to model the process which has driven the change /φ/ → /h/, as simple as possible, without loss of generality. Accordingly, we have made the following simplification: We will (i) limit our analysis to the changes in the AP of fricatives; (ii) study the change in only one (any) of the words showed in table 2; (iii) suppose there are two PS in the LS, one derived from Latin, which favors front and middle fricatives production, and other derived from prehispanic tribes, which favor middle and back production; (iv) propose there are only three possible AP in the vocal cavity, a front place (bilabial, labiodental), a middle place (dental, alveolar and post-alveolar) and a back place (palatal, velar, uvular and glottal).

In this frame, we define the main elements of the model as follows:

- $A, B$ are two groups of agents in a stratum $LS$;
- $N_A, N_B$ are the number of agents in $A, B$ and $N = N_A + N_B$ the total number of agents in $LS$;
- $S$ is the state of an agent in the $LS$ at time $t$, where $S \in \{1, 2, 3\}$, represents the AP of agent $i$, such that: $1 = \text{front}$, $2 = \text{middle}$, $3 = \text{back}$;
- $PS^A, PS^B$ are the phonological systems of $A, B$, such that $PS^A$ favors front-middle places, and conversely, $PS^B$ favors middle-back places.

In the evolutionary process, at time $t$, we randomly take from the set $A \cup B$ an active agent and a reference agent. The state of the former will change
according to the state of the latter, guided by the following imitation rules which we summarize by using the usual chemical reactions notation,

\[
\begin{align*}
A_2 + B_3 & \xrightarrow{q} A_3 + B_3 & (1a) \\
B_2 + A_1 & \xrightarrow{q} B_1 + A_1 & (1b) \\
A_3 + A_2 & \xrightarrow{p} 2 A_2 & (1c) \\
B_1 + B_2 & \xrightarrow{p} 2 B_2 & (1d) \\
A_1 & \xrightleftharpoons{r} A_2 & (1e) \\
B_2 & \xrightleftharpoons{r} B_3 & (1f)
\end{align*}
\]

For example, Eq. (1c) states that a member of the population \(A_2\) can interact with a member of \(A_3\) and that the result of the interaction are two members of the population \(A_2\).

Regarding the rules, notice that since \(PS^A\) favors front-middle and \(PS^B\) favors middle-back productions. Rules \((1a) - (1d)\) are introduced in order to emulate a process of imitation, where interactions between agents of different groups lead to non preferential states of pronunciation, and interactions between agents of the same group reinforce the preferential states of the group. In this respect, probabilities \(q, p\) define the interaction strength between agents of different groups, and the interaction strength between agent of the same group, respectively. Moreover, the noisy component expressed by rules \((1e)\) and \((1f)\), captures the variations caused by both random phonetic changes and the production of possibles allophones.

On the other hand, note that in the frame of the proposed imitation rules, the changes in the states occur only when the AP distance between the referent and the active agents are minimal. The idea here is modeling the changes in the context of close or similar sounds \([28, 27, 26]\). In this theoretical frame, our proposal has been inspired in the models of opinion formation dynamics \([8, 33]\) where social interactions, in the context of a social debate, drives the population to emergent states of consensus or polarization \([12, 10, 11]\).

Macroscopically, in the frame of evolutionary dynamics, the global state of the system can be analyzed by counting the number of agents in \(A, B\), in the states \(S = 1, 2, 3\). In the following section, based on this idea we introduce the master equation of the process and derive the evolution equations for the first moments, or mean-field approximation.

3. The equations of evolution

Let \(N_i^A, N_i^B\) be the number of agents in \(A, B\) in state \(S = i\), with \(i = 1, 2, 3\) (hereafter referred to as the occupation numbers). The master equation of the system is given by

\[
\frac{\partial P(\vec{x}; t)}{\partial t} = \sum_{\vec{y} \neq \vec{x}} T(\vec{x} | \vec{y}) P(\vec{y}; t) - \sum_{\vec{y} \neq \vec{x}} T(\vec{y} | \vec{x}) P(\vec{x}; t),
\]

where \(\vec{x} = (N_1^A, N_2^A, N_3^A; N_1^B, N_2^B, N_3^B)\) is the so called occupation vector; \(P(\vec{x}; t)\) is the probability to find the system with an occupation vector \(\vec{x}\) at
time $t$, and $T(\vec{x}|\vec{y})$ is a transition probability from a global state given by an occupation vector. In this approach, we are considering a fixed population in both groups, therefore for all $t$ we have,

$$
N_A = N_A^A + N_A^B + N_A^3,
N_B = N_B^1 + N_B^2 + N_B^3,
N = N_A + N_B.
$$

(3)

From the master equation (2), we can derive the evolutionary equations for the first moment of the occupation numbers (mean-field approximation). For instance for $N_A^2$ we have,

$$
\frac{d}{dt} \langle N_A^2 \rangle = \langle T(N_A^2 + 1|N_A^2) \rangle - \langle T(N_A^2 - 1|N_A^2) \rangle,
$$

(4)

the transitions $T$ are defined by the rules proposed in the last section, and depend on probabilities $p$, $q$, and $r$. For the case by which the system increases one agent in $N_A^2$, $T$ can be written as

$$
T(N_A^2 + 1|N_A^2) = \frac{N_A^A}{N} N_r + \frac{N_A^3}{N} \frac{N_A^A}{N - 1} p,
$$

(5)

where the first term is the probability of finding an agent of group $A$ in state 1, times the probability it randomly goes to state 2; and the second term is the probability of one interaction between an active agent of group $A$ in state $S_A^a = 3$ and a reference agent of group $B$ in state $S_B^r = 2$, leads to the former to imitate the latter with probability $p$.

For the case in which the occupation number $N_A^2$ decrease in one agent, the transition is given by,

$$
T(N_A^2 - 1|N_A^2) = \frac{N_A^2}{N} \frac{N_B^3}{N - 1} q + \frac{N_A^A}{N} r,
$$

(6)

where the first term shows the loss of an active agent of group $A$ in state 2 due to the interaction with a reference agent of group $B$ in state 3, and the second term stands for the random loss of an agent in 2 who moves to 1.

Replacing 5 and 6 in equation 4, we obtain,

$$
\frac{d}{dt} \langle N_A^2 \rangle = \frac{N_A^A}{N} N_r + \frac{N_A^3}{N} \frac{N_A^A}{N - 1} q + \frac{N_A^A}{N} r - \frac{q}{N(N - 1)} \langle N_A^3 N_B^2 \rangle - \frac{r}{N} \langle N_A^2 \rangle.
$$

(7)

From now, we will consider the evolution of the numbers are uncorrelated, then $\langle N_2^A N_2^B \rangle = \langle N_2^A \rangle \langle N_2^B \rangle$ and $\langle N_3^A N_2^A \rangle = \langle N_3^A \rangle \langle N_2^A \rangle$. Additionally, we will define, (i) $\omega_p = \frac{p}{N - 1}$, $\omega_q = \frac{q}{N(N - 1)}$, the rates of interaction intra- and inter-group; (ii) $\omega_r = \frac{r}{N}$, the rate of random changes 1 $\rightarrow$ 2; and (iii) the occupation number as fractions of the total population: $a_1 = \frac{\langle N_A^1 \rangle}{N}$, $a_2 = \frac{\langle N_A^2 \rangle}{N}$, $a_3 = \frac{\langle N_A^3 \rangle}{N}$.
\( \langle N_A^t \rangle, b_3 = \langle N_B^t \rangle \). Using the notation and approximations introduced above, we can write equation (7) as follows,

\[
\dot{a}_2 = \omega_r a_1 + \omega_p a_2 a_3 + \omega_p a_2 b_3 - \omega_r a_2
\]  

(8)

Similarly, it is possible to obtain the equations for the evolution of the other occupation numbers, which define the following system of coupled differential equations,

\[
\begin{align*}
\dot{a}_1 &= \omega_r (a_2 - a_1) \\
\dot{a}_2 &= -\omega_r (a_2 - a_1) + \omega_p a_2 a_3 - \omega_q a_2 b_3 \\
\dot{a}_3 &= -\omega_p a_2 a_3 + \omega_q a_2 b_3 \\
\dot{b}_1 &= -\omega_p b_1 b_2 + \omega_q b_2 a_1 \\
\dot{b}_2 &= \omega_p b_1 b_2 - \omega_q b_2 a_1 - \omega_r (b_2 - b_3) \\
\dot{b}_3 &= \omega_r (b_2 - b_3)
\end{align*}
\]  

(9)

It is important to highlight that since we are now working with the fractions of the occupation numbers, the constraints related to the fixed population become:

\[
\begin{align*}
n_A &= a_1 + a_2 + a_3, \\
n_B &= b_1 + b_2 + b_3,
\end{align*}
\]

and

\[
n_A + n_B = 1,
\]

where \( n_A = \frac{N_A}{N} \), \( n_B = \frac{N_B}{N} \). The constraints also show that it is possible to reduce the rank of the system to four, however, for the sake of clarity, we decided to keep the six equations for a more detailed analysis.

Lastly, note that, if there are equilibrium points in system (9), in the frame of our model, these will be related to the population reaching consensus about the manner of pronunciation. In the following section we will probe the existence of four equilibrium points, and will show that the stability of these emergent states strongly depends on the interactions rates intra- and inter-group.

4. Stability analysis

At time \( t \to \infty \), if equilibrium exists, it must satisfy \( \dot{a}_i = 0 \) and \( \dot{b}_i = 0 \). In these conditions, from Eqs. (9) and using the population constraints, it is possible to probe the existence of four non trivial equilibrium points,

I \( \vec{n}_{eq} = (0, 0, n_A; n_B, 0, 0) \)

II \( \vec{n}_{eq} = (0, 0, n_A: 0, \frac{a_B}{2}, \frac{a_B}{2}) \)

III \( \vec{n}_{eq} = (\frac{a_A}{2}, \frac{a_A}{2}, 0; n_B, 0, 0) \)

IV \( \vec{n}_{eq} = (\alpha, \alpha, Q\beta; Q\alpha, \beta, \beta) \), where \( \alpha = \frac{2n_A - Qn_B}{4 - Q^2}, \beta = \frac{2n_B - Qn_A}{4 - Q^2}, Q = \frac{\omega_q}{\omega_p} \).

The first equilibrium can only be reached if the initial conditions are set in this point, later we will show this equilibrium is unstable. Equilibria II and III are the cases where the system loses the back and the frontal AP, respectively. Equilibrium IV, on the other hand, shows a mixed final state, where the system reaches a balance among the three states of pronunciation.

The stability analysis around the equilibrium can be performed by analysing the eigenvalues of system (9). Let \( \vec{n}_{eq} = (a_1^*, a_2^*, a_3^*; b_1^*, b_2^*, b_3^*) \) be a generic equilibrium point, then,
is the Jacobian matrix evaluated in the equilibrium. Let \( \lambda \in \mathbb{C} \), an eigenvalue of \( J(\vec{n}_{eq}) \), then \( |J(\vec{n}_{eq}) - \lambda I| = 0 \). The system will have six eigenvalues; although, since there are two constrains, only four will be in general different of zero.

We are particularly interested in studying the stability as a function of \( Q \) since this parameter controls the interaction between PS\(^A\) and PS\(^B\), and also rules the population state of the equilibrium point IV. To this purpose, we set an equal population \( n_A = n_B \), and fixed the rate \( \omega_r \), such that \( r = 0.5 \), in order to focus only in the study of the effect of \( Q \) on the equilibrium conditions of the system. In this frame, we have numerically calculated the eigenvalues as a function of \( Q \), and studied the sign of the real part of the eigenvalues to evaluate the stability conditions.

The plots in figure [1] show the curves for the real part of the highest eigenvalue, \( Re(\lambda_1) \) vs. \( Q \) for the four equilibrium points. Panel A in this figure shows the calculation for state I. We can see that for all \( Q \) there is an eigenvalue with positive real part, which means this equilibrium state is always unstable. In panels B and C we can observe that states II and III behave similarly each other. This is expected because these two states are symmetric. States II and III are unstable for \( Q < 2 \) since they have at least one eigenvalue with positive real part. Finally, panel D shows the calculation for equilibrium state IV; in this case, complementary to states II and III, the equilibrium is unstable for \( Q > 2 \).

Clearly, coefficient \( Q \), which measure the relative intra- and inter-groups imitation rates, determines the stability of the different equilibrium states of the system. This is summarized schematically in figure [2] where we have highlighted possible changes of stability at \( Q = 2 \); state IV becomes unstable but we cannot ensure states II and III are stable from our stability analysis results. We can understand the changes of stability by reasoning as follows: At \( Q = 2 \), the rate of inter-group interactions equals the sum of the interaction rates within each group (\( Q = 2 \rightarrow \omega_q = \omega_{q1} + \omega_{q2} \)), and equilibrium IV becomes \( \vec{n}_{eq} = (\alpha, \alpha, Q; Q; \beta) = \left( \frac{1}{7}, \frac{1}{7}, \frac{1}{3}, \frac{1}{3}, \frac{1}{6}, \frac{1}{6} \right) \), i.e. there is a balance between the number of agents in preferential and non-preferential states of pronunciation in both groups (\( n_A^4 = n_A^2 = n_B^A \) and \( n_B^B = n_B^F \)). For \( Q > 2 \), the system loses the balance, hence equilibrium IV becomes unstable and depending on the initial conditions the system evolves to equilibrium II or III. Furthermore, in the stochastic version of the model and balanced initial conditions, fluctuations can determine the final equilibrium state, which occurs with equal probability for states II and III.

In order to visualize the time evolution of the occupation number in the system, and to study the convergence to the equilibrium points, we solved numerically the set of coupled differential equations given in Eqs. [9], and also performed numerical simulations of the stochastic agent based model, using the rules proposed in section [2] in order to test the effect of fluctuations in the dynamics of the system. For all the cases in the present study we have used the following, (i) the initial conditions were set on \( \vec{n}(t = 0) = (\frac{n_A}{2}, \frac{n_A}{2}, 0; 0, \frac{n_B}{2}, \frac{n_B}{2}) \)
(preferential states for both groups), (ii) the number of agents such that \( n_A = n_B \) and (iii) the noise \( r = 0.5 \). Figure 3 shows the evolution of the occupation numbers for \( Q = 0.7 \) in which state IV is stable. Panels A and B show the evolution in the mean field approximation (Eqs. (9)), whereas panels C and D show the outcome of the agent based model. We can see that, as expected, the system evolves toward equilibrium IV and the mean reaches the value \( \vec{n}_{eq} = (\alpha, \alpha, Q\beta; Q\alpha, \beta, \beta) = (\frac{5}{27}, \frac{5}{27}, \frac{7}{54}; \frac{7}{54}, \frac{5}{27}, \frac{5}{27}) \).

Figure 4, by contrast, shows the evolution for \( Q = 5 \) where equilibrium IV is unstable, therefore, the system should go to equilibrium II or III since we suppose they are stable; in this case, the realization showed in the plots went to II. Notably, at the beginning of the evolution, the system seems stabilize in equilibrium IV, but a larger times moves toward equilibrium II, as expected.

Clearly, the dynamics of the system seems to be ruled by several time scales. In this regards, it is well known that in the evolutionary process the time scales are related to the eigenvalues of the Jacobian. In table 3 we show the eigenvalues (calculated numerically) and its related characteristic times \( (\tau_i) \), for the case of equilibrium IV and \( Q = 5 \). See also the arrows in figure 4 panel A, which indicates the relaxation times that operate in each regime.

Additionally, we measured the distribution of times \( (T) \) needed for the system to reach the final equilibrium by performing two thousand realizations of the stochastic model. Figure 3 panel A, shows the distribution obtained. We can see a non symmetric distribution with a peak around \( T \sim 6 \times 10^5 \) and a tail at the right of the distribution. As we said before, in every realization the system stay in the unstable equilibrium IV before it reach the final stable state, therefore the total time \( T \) depends on the magnitude of the fluctuations which drive the system from the unstable equilibrium to the stable one. The latter explains the tail at the right of the distribution and also the differences in the final times observed among the stochastic simulations and the mean field approach. From the same set of two thousand realizations, we have calculated the average curve of the occupation numbers \( \left \langle \frac{N_A}{N} \right \rangle \) and \( \left \langle \frac{N_B}{N} \right \rangle \) which we show in panel B of figure 5. As compared to the mean field curves —\( a_3 \) and \( b_1 \) in Panels A and B of Fig. 4, respectively— these curves are rounded which indicate that due to fluctuations the system remains less time in state IV and also, for the same reason, the system reach rapidly the stable state.

Lastly, in order to provide a global analysis, we test the system behaviour under changes in the initial conditions, and for different values of parameter \( r \) (noise). Figure 6 shows trajectories in the plane \( a_1 - a_3 \) for \( Q = 5 \), obtained from the mean field approach simulations, where we have try different initial conditions for the numbers \( a_i \) (panel A) keeping \( r \), and where we have only changed the value of \( r \) (panel B) using the same initial condition. In panel A, we can see that depending on the proximity to the stable attractor II the system will explore the unstable equilibrium IV, as in the case of trajectories (i) and (ii), or it will not as in the case of trajectory (iii). For the case of (i) we have additionally plotted the trajectory obtained from a single stochastic realization. The plot in panel B, on the other hand, shows the trajectories seem to depend on the level of noise. The parameter \( r \) regulates the velocity of the transitions \( A_1 \leftrightarrow A_2 \) (and also \( B_2 \leftrightarrow B_3 \)). When \( r \) is strong enough, these transitions occur much faster than the rest and thus, populations \( A_1 \) and \( A_2 \) tend to have the same number of individuals, independently on the initial conditions. This
aspect explains why all the trajectories in Figure 6A rapidly approximate to the line \( a_3 = -2a_1 \). If the system starts from an initial condition with a low value of \( a_3 \), then the trajectory will be forced to go through equilibrium IV, as it can be seen from the figure. When the value of \( r \) is diminished, it is harder for the system to equal the values of \( a_1 \) and \( a_2 \) and thus, the trajectories in general deviate from the line and equilibrium IV is avoided.

In the frame of the proposed model, state II can be related to the change \(/	ext{f}/ \rightarrow /h/\) in Castilian, and state III to the observed in other Romance languages like Portuguese or Catalan. Therefore, for \( Q > 2 \), the model seems to capture very well the current pronunciations that emerged—from the real LS—in the Iberian peninsula. On the other hand, equilibrium IV, describes an emergent state of mixed pronunciation, which means there are agents using different AP to pronounce the same word. This is not observed in the real case, but can be used to understand the existence of some regionalism or local accents in the peninsula [25].

5. Conclusions

In this work, we have proposed a model to study phonetic changes in the AP for a single word as an evolutionary process guided by social interaction of imitation between two groups of people with different phonological systems. Inspired by the case of the change \(/	ext{f}/ \rightarrow /h/\) in Castilian language, we have studied a fixed population made up of two groups of interacting people, A and B, such that group A have a trend to produce frontal fricatives and conversely group B have a trend to produce back fricatives.

The rules of the model were proposed based on empirical observations, and were thought to link the phonetic changes with a process of social interactions inter- and intra-group. The model was mathematically formalized in section 2 as a stochastic process where the variable \( S \in \{1, 2, 3\} \), representing the AP for every agent in the population, changes according to the proposed interaction rules.

In this frame we studied the temporal evolution of the occupation numbers, and from first principles, we derived the coupled system of differential equations which defines the dynamics in the mean field approximation. In the equilibrium, we found three non trivial final states, which we have related to the emergence of general states of consensus in the way a word is pronounced. In this regards, we found that when the rate of interaction among agents from different groups becomes larger than the sum of the rates within each group (\( \omega_q > 2\omega_p \)), the system exhibits two emergent states (equilibria II and III) which capture very well both, the middle-back pronunciation used in Castilian and the front-middle pronunciation observed in other Romance languages. From a social point of view, we can link the condition \( \omega_q > 2\omega_p \) to the situation where the relation among individuals from different groups is larger enough to allow a common general consensus in spite of the cultural differences, in this case languages [2, 9, 31].

Lastly, it is important to say that, since social interaction are better modeled in the frame of network theory, it is necessary a complementary approach to analyze the effect of structured populations in the system. In this regards, we let an open problem to be faced in a future work.
Table 1: Phonetic consonants of the Castilian language [24], codified in the international phonetic alphabet (IPA). Rows in the table stand for the manner of articulation, and columns stand for the place of articulation. The columns have been split in three main columns representing from left to right the front, middle and back sectors in the vocal cavity. Code references: (BL) Bilabial, (LD) Labio-dental, (DE) Dental, (AL) Alveolar, (PA) Post-Alveolar, (PL) Palatar, (VL) Velar, (UV) Uvular, and (GL) Glottal. Symbols inside parenthesis (·), only appear as the result of phonetic anticipatory assimilation [27]. (*) stands for sounds not observed in all the varieties of the language.

| Place → | BL | LD | DE | AL | PA | PL | VL | UV | GL |
|---|---|---|---|---|---|---|---|---|---|
| Nasal | m | n | n | n | n | n |
| Oclusive | p, b | t, d | k, g |
| Africative | (f) | θ, (s) | f, (v) |
| Fricative | (β) | s, (z) | s, (z) |
| Approximant | | l | l |
| Place | Bilabial | Labio-dental | Dental | Alveolar | Post-Alveolar | Palatar | Velar | Uvular | Glottal |

Table 2: Some examples of Latin root words which show the change /φ/ → /h/ in Castilian (column 1). In column 2, we show the translation into Portuguese in order to show the change in this case did not happen. In column 3, we show the translation into English as a reference.

| Latin | Castilian | Portuguese | English Translation |
|---|---|---|---|
| facere | hacer | fazer | to do/ to make |
| femina | hembra | femea | female |
| ferra | hierro | ferro | iron |
| filia | hijo | filho | son |
| folia | hoja | folhagem | leaf |
| funu | humo | fumaça | smoke |

Table 3: Eigenvalues and time scales for the case \( Q = 5 \) and equilibrium IV. The time scales \( \tau_i \) in column 3 are related to the evolution showed in figure 4.

| Order \((i)\) | Eigenvalue \((\lambda_i)\) | Time scale \((\tau_i = 1/\lambda_i)\) |
|---|---|---|
| 1 | 1.0 \(10^{-5}\) | \(\sim 10^9\) |
| 2 | 0.0 | – |
| 3 | 0.0 | – |
| 4 | \(-2.5 \times 10^{-5}\) | \(\sim 10^5\) |
| 5 | \(-0.9 \times 10^{-3}\) | \(\sim 10^3\) |
| 6 | \(-1.0 \times 10^{-3}\) | \(\sim 10^3\) |
Figure 1: Stability analysis. Real part of the larger eigenvalue of the Jacobian Matrix ($\lambda_1$), as a function of the parameter $Q$. Panels A, B, C and D show the numerical calculation for the equilibria I, II, III and IV, respectively. Note that when $\text{Re}(\lambda_1) > 0$ the system is unstable, therefore equilibrium I will be unstable for all value of $Q$; II and III will be unstable for $Q < 2$, and IV for $Q > 2$.

Figure 2: The scheme shows the stability changes in the four equilibria as a function of $Q$. Solid lines represent the equilibrium in this regions is stable, and dashed lines the equilibrium is unstable.
Figure 3: The case of equilibrium IV stable. For $Q = 0.7$, panels A and B show the first moment evolution of the numbers related to group A ($a_1$, $a_2$, $a_3$) and B ($b_1$, $b_2$, $b_3$), respectively. The curves are the outcome of the evolution of Eqs. (9), which were solved by performing a Runge Kutta 8th order method. Panels C and D show the evolution of the numbers for a single realization in the agent based model. In these conditions, the population evolves to equilibrium IV, a mixed state of pronunciation where the system reaches $\bar{n}_{eq} = (\frac{5}{27}, \frac{5}{27}, \frac{7}{27}, \frac{5}{27}, \frac{5}{27}) \sim (0.18, 0.18, 0.13; 0.13, 0.18, 0.18)$.
Figure 4: The case of equilibrium II stable. For $Q = 5$, panels A and B show the first moment evolution of the numbers related to group A and B, respectively. Panels C and D, on the other hand, show the evolution of the numbers for a single realization in the agent based model. Note how at the beginning the system seems to evolve toward the unstable equilibrium IV, but escape toward the equilibrium II, which is stable in these conditions ($\vec{n}_{eq} = (0, 0, \frac{1}{2}; 0, 1, \frac{1}{4})$). The arrows in panel A, link the plot with the time scales obtained from the eigenvalues analysis (see table 3 column 3).

Figure 5: Panel A shows the distribution of final times $T$, needed to the system reach equilibrium II for $Q = 5$. Panel B, the average curves for the evolution of the numbers $a_3$ and $b_1$, averaged over 2 thousand realizations of the agent based model for $Q = 5$. 


Figure 6: Trajectories in the plane $a_1 - a_3$. Panel A, changing the initial conditions, Panel B, changing parameter $r$, for trajectory starting from initial condition than (ii) in panel A. The blue lines correspond to the solution of the differential equations, whilst the cyan curve in Panel A corresponds to a single realization of the stochastic simulation, starting from the initial condition (i).
References

[1] Martin Ball and Nicole Muller. *The celtic languages*. Routledge, 2012.

[2] Andrea Baronchelli, Maddalena Felici, Vittorio Loreto, Emanuele Caglioti, and Luc Steels. Sharp transition towards shared vocabularies in multi-agent systems. *Journal of Statistical Mechanics: Theory and Experiment*, 2006(06):P06014, 2006.

[3] Peter W Battaglia, Jessica B Hanrick, Victor Bapst, Alvaro Sanchez-Gonzalez, Vinicius Zambaldi, Mateusz Malinowski, Andrea Tacchetti, David Raposo, Adam Santoro, Ryan Faulkner, et al. Relational inductive biases, deep learning, and graph networks. *arXiv preprint arXiv:1806.01261*, 2018.

[4] Patrice Speeter Beddor. A coarticulatory path to sound change. *Language*, pages 785–821, 2009.

[5] Patrice Speeter Beddor, Anthony Brasher, and Chandan Narayan. Applying perceptual methods to the study of phonetic variation and sound change. Oxford: Oxford University Press, 2007.

[6] John D Bengtson. The basque language: history and origin. *International Journal of Modern Anthropology*, 1(4):43–59, 2011.

[7] Juliette Blevins. *Evolutionary phonology: The emergence of sound patterns*. Cambridge University Press, 2004.

[8] Claudio Castellano, Santo Fortunato, and Vittorio Loreto. Statistical physics of social dynamics. *Reviews of modern physics*, 81(2):591, 2009.

[9] Xavier Castelló, Víctor M Eguíluz, Maxi San Miguel, Lucía Loureiro-Porto, Rüttta Toivonen, J Saramäki, and K Kaski. Modelling language competition: bilingualism and complex social networks. In *The evolution of language*, pages 59–66. World Scientific, 2008.

[10] Andrés Chacoma, Germán Mato, and Marcelo N Kuperman. Dynamical and topological aspects of consensus formation in complex networks. *Physica A: Statistical Mechanics and its Applications*, 495:152–161, 2018.

[11] Andrés Chacoma and Damián H Zanette. Critical phenomena in the spreading of opinion consensus and disagreement. *arXiv preprint arXiv:1412.6823*, 2014.

[12] Andres Chacoma and Damian H Zanette. Opinion formation by social influence: From experiments to modeling. *PloS one*, 10(10):e0140406, 2015.

[13] Imen Daly, Zied Hajaiej, and Ali Gharsallah. Physiology of speech/voice production. *Journal of Pharmaceutical Research International*, pages 1–7, 2018.

[14] W Tecumseh Fitch. The biology and evolution of speech: a comparative analysis. *Annual Review of Linguistics*, 4:255–279, 2018.
[15] James E Flege, Carlo Schirru, and Ian RA MacKay. Interaction between the native and second language phonetic subsystems. *Speech communication*, 40(4):467–491, 2003.

[16] Paul Foulkes. Historical laboratory phonology—investigating /p/ > /f/ > /h/ changes. *Language and Speech*, 40(3):249–276, 1997.

[17] Michael Gregory and Susanne Carroll. *Language and situation: Language varieties and their social contexts*. Routledge, 2018.

[18] Jonathan Harrington, Felicitas Kleber, Ulrich Reubold, Florian Schiel, Mary Stevens, and P Assmann. The phonetic basis of the origin and spread of sound change., 2019.

[19] Charles F Hockett. Linguistic elements and their relations. *Language*, 37(1):29–53, 1961.

[20] Peter Ladefoged and Sandra Ferrari Disner. *Vowels and consonants*. John Wiley & Sons, 2012.

[21] Erez Lieberman, Jean-Baptiste Michel, Joe Jackson, Tina Tang, and Martin A Nowak. Quantifying the evolutionary dynamics of language. *Nature*, 449(7163):713, 2007.

[22] Philip Liebermann. *The speech of primates*, volume 148. Walter de Gruyter GmbH & Co KG, 2019.

[23] Paul M Lloyd. *From Latin to Spanish: Historical phonology and morphology of the Spanish language*, volume 173. American Philosophical Society, 1987.

[24] Eugenio Martínez-Celdrán, Ana Ma Fernández-Planas, and Josefina Carrera-Sabaté. Castilian spanish. *Journal of the International Phonetic Association*, 33(2):255–259, 2003.

[25] Fernando Martínez-Gil et al. *Issues in the phonology and morphology of the major Iberian languages*. Georgetown University Press, 1997.

[26] Kuniko Y Nielsen. Implicit phonetic imitation is constrained by phonemic contrast. In *Proceedings of the 16th International Congress of the Phonetic Sciences, Saarbrücken, Germany, Saarbrücken, Germany: The International Congress of the Phonetic Sciences*, volume 1964. Citeseer, 1961.

[27] John J Ohala et al. The phonetics and phonology of aspects of assimilation. *Papers in laboratory phonology*, 1:258–275, 1990.

[28] Jennifer S Pardo. On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America*, 119(4):2382–2393, 2006.

[29] Gregorio Salvador. Hipótesis geológica sobre la evolución f -> h. *Introducción Plural a la Gramática Histórica (F. Marcos Marín, ed.), Madrid, Cincel*, pages 11–21, 1983.

[30] Caroline Smith. Handbook of the international phonetic association: A guide to the use of the international phonetic alphabet (1999). *Phonology*, 17:291–295, 08 2000.
[31] Dietrich Stauffer and Christian Schulze. Microscopic and macroscopic simulation of competition between languages. *Physics of Life Reviews*, 2(2):89–116, 2005.

[32] Pengfei Sun, David A Moses, and Edward F Chang. Modeling neural dynamics during speech production using a state space variational autoencoder. In *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 428–432. IEEE, 2019.

[33] György Szabó and Attila Szolnoki. Three-state cyclic voter model extended with potts energy. *Physical Review E*, 65(3):036115, 2002.

[34] Stephanie R Thorpe. A guide to the international phonetic alphabet and its application. *The Choral Journal*, 59(9):51–59, 2019.

[35] Gracian Trivino and Michio Sugeno. Towards linguistic descriptions of phenomena. *International Journal of Approximate Reasoning*, 54(1):22–34, 2013.

[36] Grant Walker and Gregory Hickok. Speech production. In *The Oxford Handbook of Psycholinguistics*, page 291. Oxford University Press, 2018.