Int. J. Appl. Math. Comput. Sci., , Vol. , No. , –
DOI:

EVOLUTIONARY LANGUAGE COMPETITION - AN AGENT-BASED MODEL

MICHAŁ GRZEJDZIAK [a]

[a]Faculty of Mathematics and Computer Science, Jagiellonian University, Łojasiewicza 6, 30-348 Kraków, Poland

In this paper, we present a novel approach to the problem of language competition, studying it in the context of language evolution. We propose a new agent-based model of this phenomenon within the framework of naming games used in evolutionary linguistics, analyze its behaviour for different combinations of parameters and compare our results with state-of-the-art model of D. M. Abrams and S. H. Strogatz. In many situations, we get the final scenario of dominance-extinction, but in some cases, our model predicts that language competition can lead to the emergence of a new mixture language. We believe that our work opens new perspectives for the study of language competition, showing that it can be analyzed as a particular example of coevolution of languages.

Keywords: language competition, language extinction, language evolution, agent-based models, language games, naming game

1. Introduction

In the era of globalization and the Internet, language extinction is a particularly escalating and difficult problem. According to the alarming UNESCO data from 2010 (Moseley, 2010), over 40% of around 6000 world languages are threatened with extinction within the next 100 years. Among the most endangered languages in the world, which can extinct with the next few decades, are for instance Karaim in Ukraine, Argobba in Ethiopia, Ava-Cenoeiro in Brazil, and Green Gelao and Red Gelao in China (Moseley, 2010). With each language extinction, the world irretrievably loses elements of its cultural heritage and diversity.

This problem was recently widely discussed among researchers (Schulze and Stauffer, 2006; Castelló *et al.*, 2013). The language death has been explained as a result of the phenomenon of language competition, where several languages in a given area compete with each other for users, and several mathematical models has been proposed to understand better this process, its causes and course (Abrams and Strogatz, 2003; Patriarca and Leppännen, 2004; Castelló *et al.*, 2008; Patriarca *et al.*, 2012; Castelló *et al.*, 2013; Wang and Minett, 2005; Wang and Minett, 2008).

In this paper, we address the problem of language competition with a new agent-based model. In contrast to the previous work, we analyze language competition within the framework of evolutionary linguistics and treat the phenomenon as a particular example of language coevolution.

1.1. Related work. Our work is closely related to two areas. The first is modelling of language competition, and the second is computer modelling of language evolution.

Many models of language competition were proposed during the past two decades. In 2003, D. M. Abrams and S. H. Strogatz proposed a model in the form of differential equation (Abrams and Strogatz, 2003). Their paper was widely discussed in further work concerning language competition (Patriarca and Leppännen, 2004; Castelló *et al.*, 2008; Patriarca *et al.*, 2012; Castelló *et al.*, 2013; Wang and Minett, 2005; Wang and Minett, 2008). More complex models were proposed by M. Patriarca and T. Leppannen (Patriarca and Leppännen, 2004) to address the role of geographical distribution of population in the process of language competition or by W. S-Y. Wang and J. W. Minett (Wang and Minett, 2005; Wang and Minett, 2008) to include the potential presence of bilinguals in the context of the phenomenon. In the series of articles, X. Castelló with collaborators (Castelló *et al.*, 2008; Patriarca *et al.*, 2012; Castelló *et al.*, 2013) summarized the previous work and translated some of the previous models to agent-based ones so that the authors could study language competition depending on population's structure modelled in the form of social networks.

Computer modelling, especially agent-based, is one of the main tools to study language evolution since James Hurford's paper from 1989 (Hurford, 1989). Starting from the 1990s, the framework of computer models based on language games in a population of agents has been developed and analyzed in the works of such researchers as L. Steels (Steels, 1995; Steels, 1996; Steels and Loetzsch, 2012), A. Baronchelli (Baronchelli *et al.*, 2006; Baronchelli *et al.*, 2008), or P. Vogt (Vogt, 2009).

To our knowledge, these two areas were rarely considered in connection with each other before, and language competition was described in terms of language shifts in a community of individuals rather than language evolution. This fact is stated most explicitly by X. Castelló *et al.* that their ”aim is to study language shift, rather than language variation and change” (Castelló *et al.*, 2013).

1.2. Motivation. In this paper, in contrast to the previous work, we study language competition in the context of language evolution, assuming that competing languages coevolve. We propose an agent-based model based on the concept of naming games (Steels, 1995; Lipowska, 2011) and show that it can give similar outcomes to state-of-the-art model of D. M. Abrams and S. H. Strogatz (Abrams and Strogatz, 2003). We believe that our results show that language competition can be studied within the framework of evolutionary linguistics.

We hope that our work will provide new perspectives for further research concerning language competition and language evolution. For instance, the evolution of Creole languages, the vernacular mixture languages that emerged in European colonies during the 17th and 18th century (Mufwene, 2017), was already analyzed with agent-based models (Jansson *et al.*, 2015), and could be potentially studied from the perspective of language competition within the framework proposed in this paper. We believe that our work could also be helpful in various problems of social sciences and humanities. The study of history knows many cases of language competition, for instance the extinction of East Germanic languages (Visigothic, Ostrogothic or Vandalic) (Moulton and Buccini, 2014) in the Middle Ages. These languages, despite their dominating status as the language of rulers, ”lost” in the language competition against Latin and Roman languages. Last but not least, we believe that our work could suggest some solutions in the design of language-preserving programmes (Abrams and Strogatz, 2003).

Addressing the above motivations, in this paper we will focus on the following questions:

- How does language competition proceed, depending on initial conditions?
- Can a new language, being a mixture of competing languages, emerge in language competition? If yes, under what conditions?
- What is the influence of language's overt prestige in language competition? Overt prestige of language can be understood as its presence in administration, religion, or education (Castelló *et al.*, 2013), e.g. Latin in the Medieval Church, official language in a state, or English in contemporary culture and science.
- What impact on language competition do some mechanisms of language preservation have?

1.3. The paper's structure. The paper is structured as follows. First, we introduce some necessary terms concerning evolutionary linguistics and language games, and state the problem of language competition formally within the evolutionary framework in the section 2. Next, we define the rules of the proposed model in the section 3, and describe the simulation process in the section 4. We present and analyze our results from several experiments in the section 5. Finally, we summarize the paper in the section 6.

2. Evolutionary linguistics framework

In this section, we provide the formal statement of language competition within the framework of evolutionary linguistics, using the definitions of units, levels, and mechanisms of language evolution provided by Nathalie Gontier (Gontier, 2017). Also, we describe briefly the concept of naming games.

2.1. The formal statement of language competition. The problem of language competition formalized on the grounds of applied evolutionary epistemology, a theory developed by Nathalie Gontier (Gontier, 2017), who emphasizes the necessity of identification of language evolution units, levels and mechanisms, which are defined as below (Gontier, 2017):

1. ”X is a unit if one can minimally point out one level where X evolves, and one mechanism whereby X evolves” (*What evolves?*).
2. ”X is a level if one can minimally point out one unit that evolves by minimally one mechanism at X” (*Where does it evolve?*).
3. ”X is a mechanism if one can minimally point out one unit that evolves at one level by means of X” (*How does it evolve?*).

We will define the language competition phenomenon as follows: in an area there is a population, in which each member at an initial moment uses one of two different languages. Then, as a result of local interactions between members of population, both languages evolve. Now, we can take the competing languages as units, the population as level and the local interactions as mechanisms of language evolution. Furthermore, we can identify the language competition phenomenon as an example of language coevolution. Comprehensive approach to the problem would demand an analysis of dynamics and correlations of all elements of language (morphology, syntactics, phonology, etc.), which is probably impossible to include in a single model. For this reason, in this paper, we will limit the problem to the coevolution of lexical systems (lexicons), assuming that other elements of language have only a negligible impact on it. Therefore, in the proposed model we will examine lexicons as units of language evolution.

2.2. Naming games. The proposed model is based on a tool frequently used in evolutionary linguistics: naming games in an agent-based model (Steels, 1995; Lipowska, 2011; Lipowska, 2016). It can be shortly described as follows: each unit of a given population is identified with a part of a computer program, which is called an agent. Each agent has a set of (object, list of words) pairs, which is called a lexicon. Agents interact pairwise in naming games, during which one agent is called a speaker, the other one is called a listener. The speaker chooses one object and communicates one of the words from the corresponding list to the listener. If the listener has this word in the corresponding list in its lexicon, we call the interaction ”successful”, otherwise we call the interaction ”unsuccessful”. Depending on the result, the agents modify their lexicons following specifically defined rules.

In the majority of naming games models, the interactions concerning different objects are independent from each other. Therefore, apart from one simulation in an n-object environment, n independent simulations in n one-object environments can be performed. However, in the context of language competition, such an assumption seems to be too simplifying. Competing languages can vary a lot in their perception of the world, i.e. the sets of objects recognized by them can be very different. We will take into account this observation in the definition of the proposed model.

3. The model's description

In this section we define the rules of the proposed model. We describe the details of agents and languages in the subsection 3.1. The rules of interactions between agents are defined in the subsection 3.2. At the end of this section, in the subsection 3.3, we define the model's additional variants.

3.1. Agents and languages. In the proposed model, there is a population (a set) of N (for simplicity $N = n^2$ for some $n \in \mathbb{N}$) agents in a d-object environment. Agents have lexicons, which are sets of (object, list of words) pairs. Words and objects are represented by integer values. Each word in a list associated with an object is assigned weight, a number from the $[0.0; 3.0]$ interval, which represents the degree of the agent's acquaintance with it. The weight of 0.0 denotes here complete unawareness of a word, weights other than 0.0 higher degrees of acquaintance. The weight of 3.0 denotes complete acquaintance with a word. Each word belongs to one of two languages (sets of words), which we denote by ℓ_1 (language 1) and ℓ_2 (language 2). A word belongs to ℓ_1 if its integer value is odd; otherwise it belongs to ℓ_2. In a lexicon of an agent X, a word of the highest weight of those in the list corresponding to an object will be called the dominating word of this object. On this basis we can specify the notion of the dominating language of the agent X. Let $dom_i(X)$ denote the total number of dominating words from ℓ_i in the X's lexicon and $sum_i(X)$ denote the sum of weights of all words from ℓ_i in the X's lexicon (for $i = 1, 2$). The following criteria will determine the dominating language of X:

- If $dom_1(X) > dom_2(X)$, then ℓ_1 is the dominating language of X.
- If $dom_1(X) < dom_2(X)$, then ℓ_2 is the dominating language of X.
- If $dom_1(X) = dom_2(X)$ and $sum_1(X) > sum_2(X)$, then ℓ_1 is the dominating language of X.
- If $dom_1(X) = dom_2(X)$ and $sum_1(X) < sum_2(X)$, then ℓ_2 is the dominating language of X.
- If none of the above occurs, then ℓ_1 is the dominating language of X (it is chosen arbitrarily, but such a situation is very unlikely in the model).

An agent of which the dominating language is ℓ_i will be called i-lingual. Let P denote a population of agents. We define the following characteristics of languages:

- The population of ℓ_i will be the set $P_i \subset P$ of all i-lingual agents. Its size will be denoted by N_i.
- The environment of ℓ_i will be the set of all objects recognized with a weight greater than 0.0 with a word belonging to ℓ_i in any agent's lexicon. Its size will be denoted by d_i. The ratio $\frac{d_i}{d}$ will be called the degree of development of ℓ_i.

Object	Word	Weight
1	1	3.0
2	3	3.0
3	5	3.0
4		

Table 1. 1-lingual agent initialization for $d = 4$, $d_1^0 = 3$

Object	Word	Weight
1	2	3.0
2	4	3.0
3		
4		

Table 2. 2-lingual agent initialization for $d = 4$, $d_2^0 = 2$

- The spread of ℓ_i will be the ratio

$$s_i = \frac{\sum_{X \in P} sum_i(X)}{3.0Nd} \tag{1}$$

In other words, it will be the ratio of the sum of weights of all words belonging to ℓ_i in any agent's lexicon to the maximal sum of weights in a monolingual population. It will be denoted by s_i.

Obviously, in any moment we have:

$$N_1 + N_2 = N \tag{2}$$

and

$$\forall_{i=1,2}\ d_i \leqslant d \tag{3}$$

To represent the geographical distribution of the population, agents are placed in the vertices of a weighted $n \times n$ lattice graph of edges of weights $1.0 + \epsilon$, complemented with edges of weights ϵ to the complete graph. Vertices are counted from the first row and the first column: vertex i is the vertex of coordinates $(\lfloor \frac{i}{n} \rfloor, i \pmod n)$. The weight w_{ij} of the edge connecting vertex i with vertex j is given by:

$$w_{ij} = \begin{cases} 1.0 + \epsilon & \text{if } i = j \pm 1 \text{ or } i = j \pm n \\ \epsilon & \text{otherwise} \end{cases} \tag{4}$$

Initialization of the agents and their lexicons is determined by the parameters N_1^0, N_2^0, d_1^0 and d_2^0. For both languages, N_i^0 agents are initialized as i-lingual, with lexicons containing one word for d_i^0 objects; for instance, an object j ($j = 1, 2, \ldots, d_i^0$) is initialized with the word of value $2 * (j - 1) + i$ and the weight of 3.0. A sample initialization of lexicons is shown in the tables 1 and 2. 1-lingual agents are initially placed in the vertices of the indices from 1 to N_1^0, 2-lingual agents are placed in the remaining vertices. A sample initialization of the structure is shown in the figure 1.

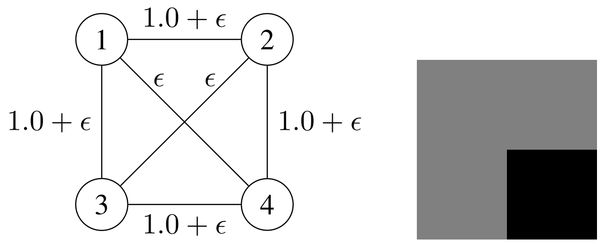


Fig. 1. Initialization of the population structure given by lattice graph for $N = 4$, $N_1^0 = 3$, $N_2^0 = 1$. 1-lingual agents are coloured grey, 2-lingual agents are coloured black

3.2. Interactions between agents. Interactions between agents are defined as follows: at first a number $x = 1, \ldots, N$ is randomly generated[1] and the agent x (i.e. agent in the vertex x) becomes a speaker. Then, a listener is randomly selected. The probability p_{xy} of selecting agent y as a listener is given by:

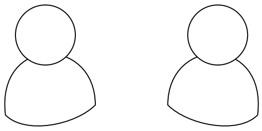
$$p_{xy} = \frac{w_{xy}}{\sum_{k=1}^{N} w_{xk}} \tag{5}$$

The speaker randomly chooses number $o = 1, \ldots, d$. If it has in its lexicon at least one word corresponding to the object o, then it communicates to the listener the dominating word of this object. Otherwise, it creates a new corresponding word in the following way: if it is i-lingual, it randomly generates an integer number k and the new word receives the value of $2k + i$ and the weight of 1.0, and is communicated to the listener. If the listener recognizes the communicated word, then the interaction is successful, otherwise it is unsuccessful. Agents keep the records of their successful interactions for both languages and on this basis they modify their lexicons after each interaction. If the interaction is successful, then both agents increase the weights of the communicated word by 0.3 multiplied by the ratio of the number of successful interactions in the corresponding language in which they participated, to the total number of interactions in this language in which they participated. They also decrease by the same value weights of all other words corresponding to the object o in their lexicons. So defined behaviour after successful interactions is reflecting the assumption, that during language competition agents prefer using more successful language. Value 0.3 is adopted conventionally; by that means it is necessary for an agent to interact in at least 10 successful interactions concerning one word to increase its weight by maximal available value 3.0. If the interaction is unsuccessful, the listener acquires the communicated word with the

[1] We generate random numbers with uniform distribution

Speaker

Object	Word	Weight
1	**1**	**2.1**
2	3	2.7
	4	0.7
3	5	3.0
4		



Communicated word: **1**

Listener

Object	Word	Weight
1	**1**	**0.4**
	2	**2.7**
2	4	2.7
3	6	1.8
4		

Lexicons before the interaction

Speaker

Object	Word	Weight
1	**1**	**2.4**
2	3	2.7
	4	0.7
3	5	3.0
4		

Listener

Object	Word	Weight
1	**1**	**0.7**
	2	**2.4**
2	4	2.7
3	6	1.8
4		

Lexicons after the interaction

Fig. 2. A sample successful interaction, for $d = 4$. For simplicity, for both agents respective ratios of successess to all interactions are assumed 1.0.

weight 1.0 and the speaker decreases its weight by 0.3. When the weight of a word decreases to 0.0 in an agent's lexicon, then it is removed unless it is the only word in this lexicon corresponding to an object; if so, it is left with a conventional weight of 0.1. It is assumed here, that the process of acquiring a word is more effective than the process of learning, and that the only word denoting an object cannot be forgotten. A sample interaction between agents is shown in the figures 2. For simplicity, for both agents respective ratios of successess to all interactions are assumed 1.0.

3.3. The model's variants. The above rules describe the basic variant of the proposed model. In addition, two other variants will be analysed: one with a "total speaker" agent, the other one with a "total listener" agent. Both variants include v additional interactions between agents from the population and respectively a total speaker or a total listener before each simulation step[2]. The total speaker has its own lexicon, which contains one word belonging to the variant language ℓ_{i_v} for $d^0_{i_v}$ objects; it plays the role of speaker in each interaction it participates in. It can reflect the influence of languages' overt prestige in the population. The total listener plays the role of listener in each interaction it participates in; each such an interaction is successful. It can reflect some mechanisms of language preservation.

[2]the k-th simulation step (or the k-th iteration) stands for N subsequently simulated interactions, from the $((k-1)N+1)$-th one to the kN-th one ($k \in \mathbb{N}_1$)

4. The model's simulations

In this section, we summarize briefly the model's parameters and describe the simulation process for the experiments presented in the section 5.

The set of paramaters of the model include:

- the n parameter determining the population's size $N = n^2$, together with the initial sizes N^0_1 and N^0_2 of both languages' populations
- the number d of all objects and the initial environments' sizes d^0_1 and d^0_2 of both languages
- the ϵ parameter determining the weights of edges between agents
- the model's variant: basic, with a total speaker or with a total listener (together with the variant language parameter i_v and the variant influence parameter v)

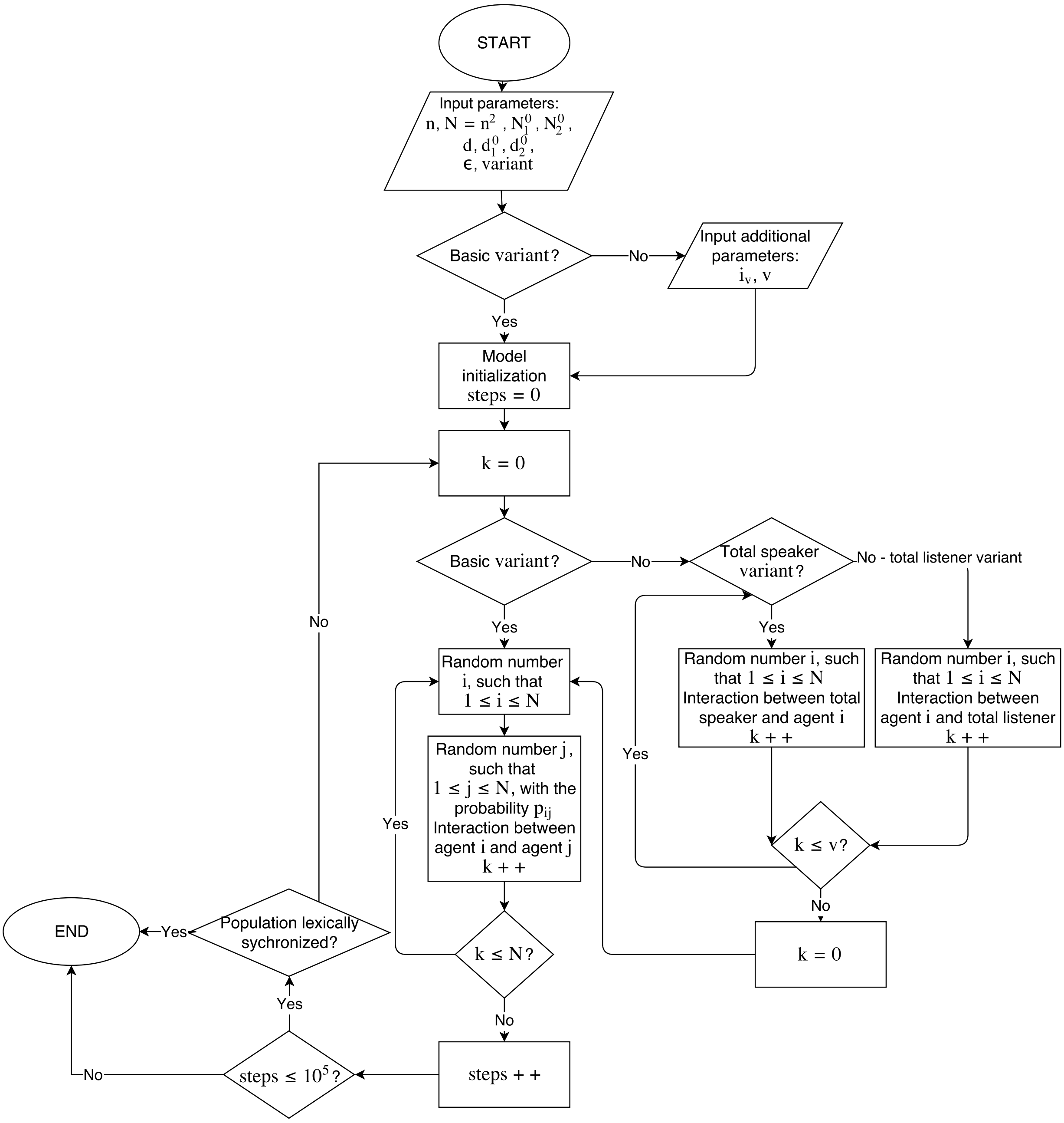


Fig. 3. The flowchart of the model's simulations

To study the model's behaviour, computer simulations were performed in a program implemented in Java. Source code is available at `https://github.com/Grzejdziok/Evolutionary-language-competition`. Each of simulations was performed in 10^5 steps but was terminated when the population of agents achieved the state of lexical synchronization (which means that agents' lexicons are identical). The figure 3 shows the flowchart of the model's simulations.

All simulations analyzed here were performed with $n = 4$ ($N = 16$), $d = 5$, $\epsilon = 0.05$. For each set of parameters, at least 1000 independent simulations were taken, and their results were averaged. During conducting the simulations, the dynamics of characteristics of both languages: their degrees of development, spreads and numbers of users were measured.

5. Experiments

In this section we present and analyse the results of simulations of the proposed model. In the subsection 5.1,

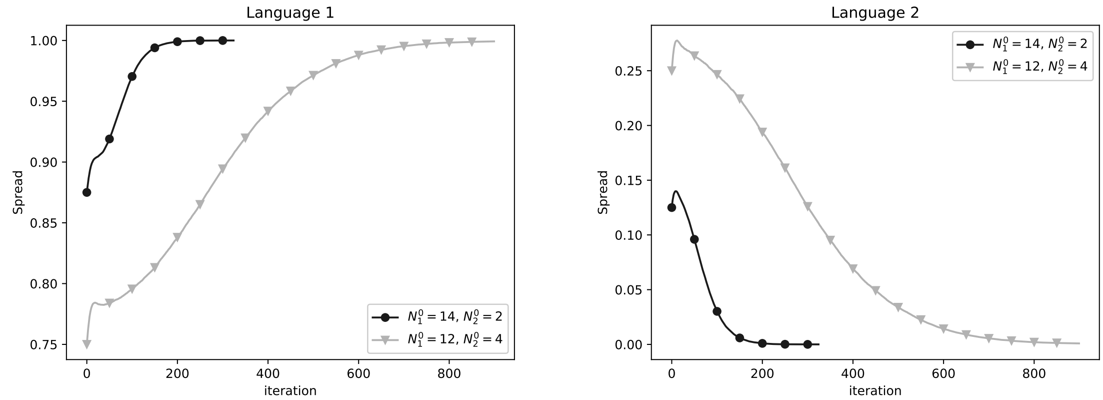


Fig. 4. The dynamics of the spreads of competing languages for different initial sizes of populations

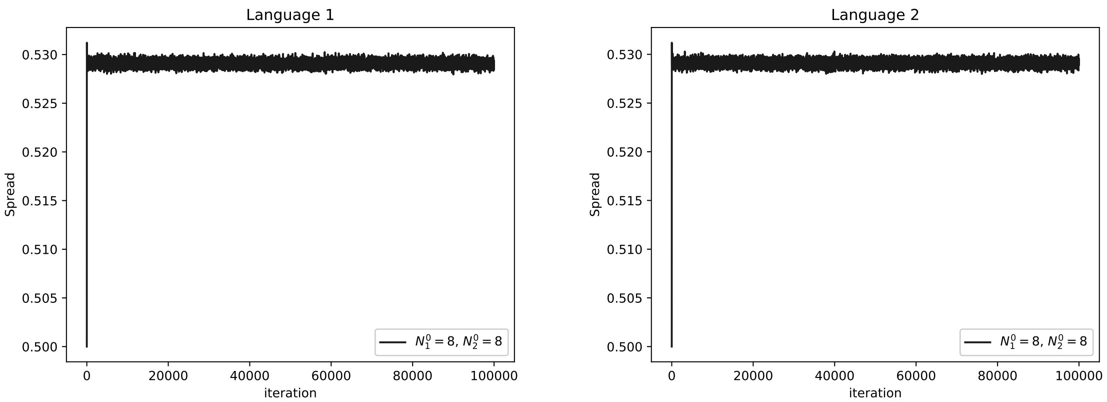


Fig. 5. The dynamics of the spreads of competing languages for equal initial sizes of populations

we study how differences in intial sizes of populations impact on the model's behaviour. In the subsection 5.2, we analyse the impact of different initial environments' sizes of the languages. In the subsection 5.3, we study joint impacts of different initial sizes of both populations and environments. In the remaining subsections 5.4 and 5.5, we analyse the model's behaviour in the presence of the model's variants.

5.1. Initial populations' sizes' impact. In order to examine the model's behaviour depending on the parameters N_1^0 and N_2^0, simulations for four different sets of parameters were performed. The dynamics of spreads of competing languages obtained is shown on the plots in the figure 4. We can see that any disproportion in initial sizes of language populations results in the total extinction of the initially outnumbered language, and the eventual common agents' lexicon consists only of the words from the language of the greater initial population. It is worth noting, that the less disproportion in initial sizes is, the longer language competition lasts: for the ratio 12.5% : 87.5% of sizes competition lasts for 151 iterations on average, for 25% : 75% it lasts for 547 iterations, and for 37.5% : 62.5% it lasts for 3.181 iterations. Therefore, although the final state is identical in all cases, the time to reach it is very sensitive to changes in initial ratio of language populations' sizes. The results obtained are similar to the actual situations, e. g. to the earlier mentioned extinction of East Germanic languages (Moulton and Buccini, 2014).

For the ratio 50% : 50% population did not reach the state of lexical synchronization in any simulation. For this case, the plot of the spreads of competing languages is shown in the figure 5. The spreads of both languages oscillate around 53%, what probably means that the languages stablize in initial populations while constantly penetrating the border between the languages' populations. The penetration is performed only to a small extent, and none of the languages can get an advantage. The population of agents reaches an equilibrium, in which three regions can be distinguished: one of ℓ_1, one of ℓ_2, and one bilingual on the border.

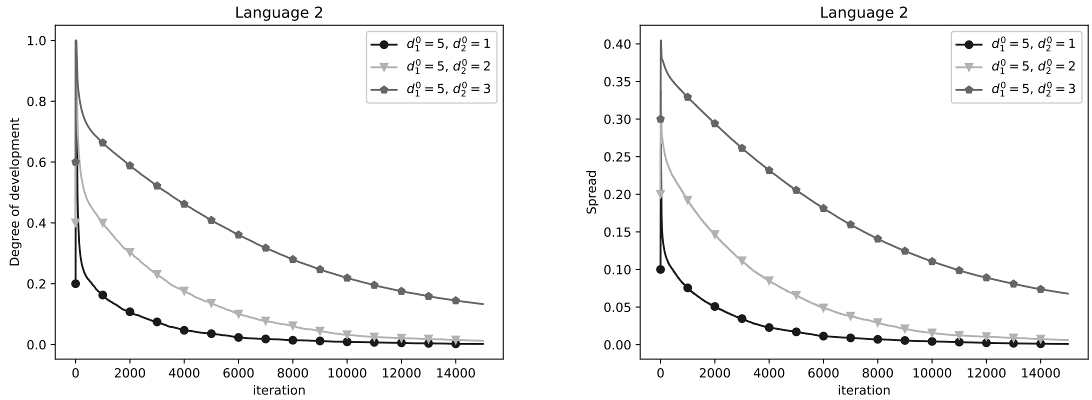


Fig. 6. The dynamics of the degree of development and the spread of the language initially less developed for simulations with equal initial populations' sizes.

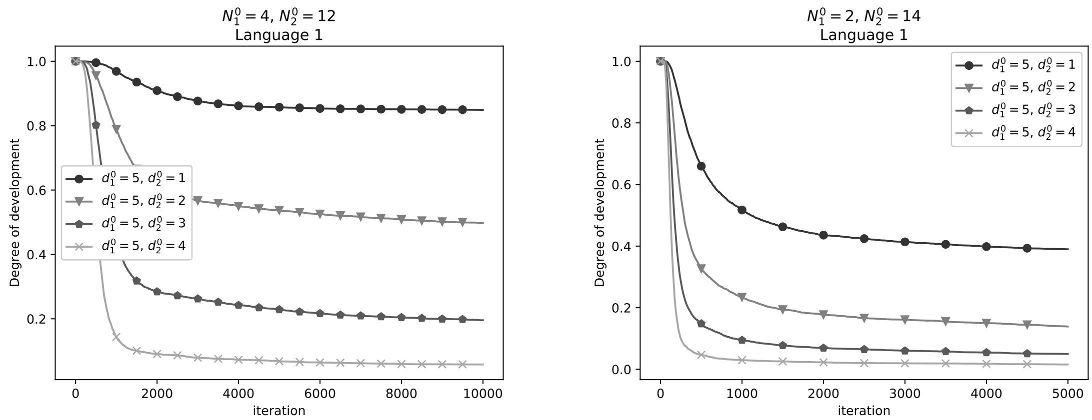


Fig. 7. The dynamics of the degrees of development of the language initially less developed for different initial population sizes and different initial environments' sizes.

5.2. Initial environments' sizes impact. The influence of initial environments' sizes d_1^0 and d_2^0 was examined in simulations with equal initial populations' sizes. The results are shown in the figure 6. The plots show that, in general, the language of the smaller initial environment's size totally extincts. Therefore, a disproportion in initial environments' sizes has similar effect to a dispropotion in initial sizes of languages' populations. An essential difference can be observed between times of reaching lexical synchronization. In the case of decreasing disproportion in initial environments' sizes - similarly to disproportion in sizes of initial language populations - an increase in time of reaching the state of lexical synchronization can be observed, and it is similarly sensitive to changes in the parameter: for d_1^0, d_2^0 equaling respectively 5 and 1, the competition lasted for 3.061 iterations on average, for 5 and 2 it lasted for 6.214 iterations, and for 5 and 3 it lasted for 16.198 iterations. A slightly different behaviour can be observed for $d_1^0 = 5$ and $d_2^0 = 4$; in 69% performed simulations the language initially less developed totally extincts after 28.751 on average, in the other 31% simulations the population reaches an equilibrium similar to this observed in the case when $N_1^0 = N_2^0$ and $d_1^0 = d_2^0$. Probably, it is a result of indeterministic time of finding a common word for an initially unknown object by the less developed language; in some cases it happens so quickly, that the other language does not manage to get an advantage. However, in most cases it happens too slowly.

5.3. Joint impact of initial populations' sizes and initial environments' sizes. Simulations performed for different sets of parameters has shown, as expected, that if one language has initially both greater population's size and environment's size, then the other one totally extincts in competition; the eventual common agents' lexicon consists only of the words from the language initially in advantage. Interesting results can be observed in a

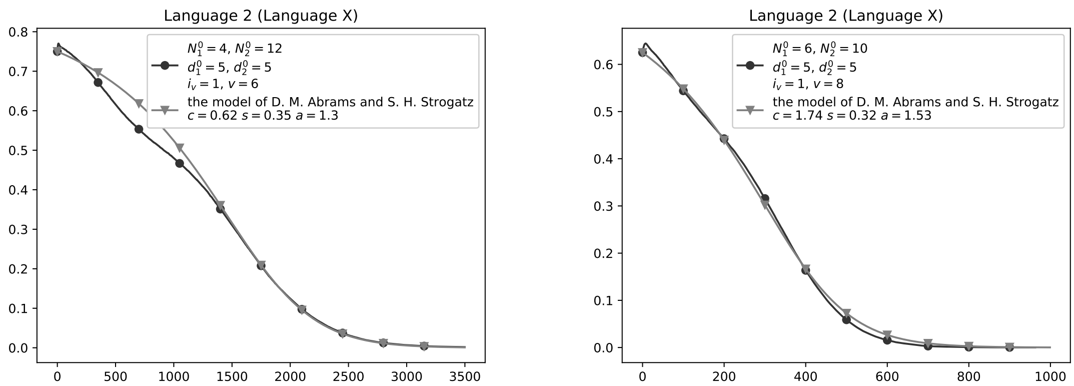


Fig. 8. A comparison of the model proposed in this paper and the model of D. M. Abrams and S. H. Strogatz. In both cases, the "total speaker" uses ℓ_1.

situation, when initially one language has an advantage in population's size, and the other in environment's size. Some examined cases are shown in the figure 7. It appears, that in the proposed model, in such situations, the eventual common agents' lexicon can be based on both languages. Such a behaviour could not appear in the previous models of language competition (Abrams and Strogatz, 2003; Patriarca and Leppännen, 2004; Castelló *et al.*, 2008; Castelló *et al.*, 2013; Wang and Minett, 2005; Wang and Minett, 2008). In the proposed model, it occurs for relatively big disproportion in populations' sizes (at least 25% : 75%) and relatively big disproportion in environments' sizes (for the ratio of population sizes 25% : 75% it is 5 : 2, for 12.5% : 87.5% it is 5 : 1). Simulations with a smaller disproportion in populations' sizes resulted in a total extinction of the initially less developed language. This phenomenon of language mixture, appearing in the model, can be observed also in reality: an example here could be language competition between Latin and European ethnic languages, which resulted in many loanwords in them. In this context, the development of Creole languages can also be considered.

5.4. "Total speaker" impact. The "total speaker" variant in the model allows to examine the impact of languages' overt prestige or, in other words, the notion of language "status". This term was considered by D. M. Abrams and S. H. Strogatz (2003) as a numerical parameter of their model in the form of the following differential equation:

$$\frac{dx}{dt} = yP_{yx}(x,s) + xP_{xy}(x,s) \quad (6)$$

where x and $y = 1 - x$ denote the percentile numbers of users of X and Y languages, and P_{yx} denotes the average probability of language change to the language X by the users of language Y. The above equation can be expressed as

$$\frac{dx}{dt} = ycx^a s + xcy^a(1-s) \quad (7)$$

where s stands for the relative "status" of the language X to the language Y, and c and a are constants. For known cases, the value of a was found roughly constant and equal to 1.31 ± 0.25 (where 0.25 stands for standard deviation). The status parameter was described as "a parameter that reflects the social or economic opportunities afforded to its speakers" (Abrams and Strogatz, 2003), and its value is adjusted to statistical data. However, this approach offers only little explanatory value; appropriate values of the status parameter can be adjusted, but it is hard to determine what they could mean. Simulations performed in the model proposed in this paper has shown that the "total speaker" variant can give similar results to those of the model proposed by D. M. Abrams and S. H. Strogatz (2003). A comparison of both models is shown in the figure 8. The similarity in behaviour of both models indicates, that the "total speaker" variant can be a good representation of the language's status. We believe that in this form, it is much easier to interpret than in the form of a numerical value.

5.5. "Total listener" impact. The "total listener" variant, which can imitate diverse language-preserving activities, was examined in experiments conducted with different values of the variant influence v. The simulations has shown, that such activities can extend the time of competition and delay the extinction of the initially outnumbered language. The results of the experiments are shown in the figure 9. The time extension, however, is relatively small: for $v = 0$ the average time of reaching the state of lexical synchronization is 547 iterations, for $v = 2$ it is 760 iterations, for $v = 4$ it is 1.073 iterations, and for $v = 6$ it is 1.506 iterations. In the last case,

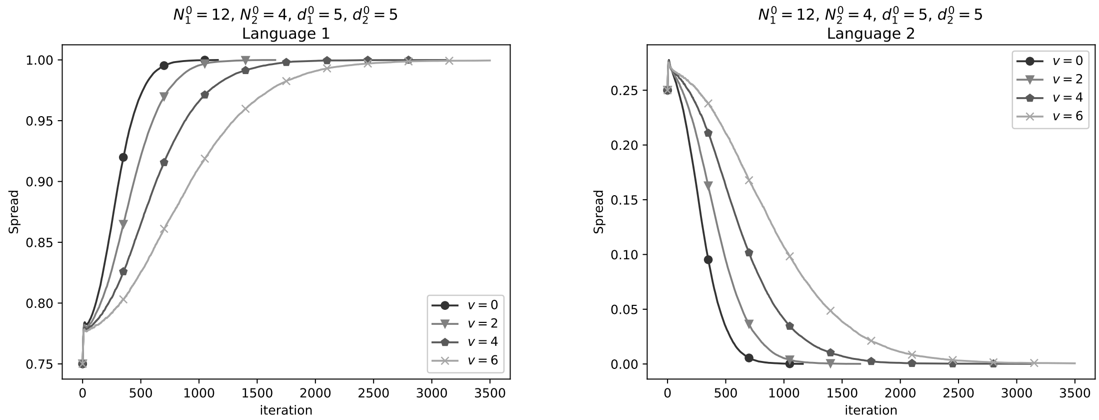


Fig. 9. The spread of languages for the simulations in the ”total listener” variant.

language competition lasts for less than three times longer than in the case of $v = 0$ on average.

6. Conclusions

In this paper, we described language competition in the context of evolutionary linguistics and proposed a new agent-based model of this phenomenon based on the concept of naming games. Similarly to the real-world data and previous models, our model predicts language death as a result of language competition in many situations. However, in contrast to the previous work, in some cases our model predicts that a new mixture language can emerge. We proposed the mechanism of ”total speaker” to model explicitly the impact of overt prestige in language competition and showed that it has similar impact as the language status parameter in state-of-the-art model of D. M. Abrams and S. H. Strogatz (Abrams and Strogatz, 2003). All in all, we believe that our results cogently show that language competition can be studied within the framework of evolutionary linguistics. It appears that local interactions in a population of agents can be sufficient to describe this phenomenon, and therefore it can be analyzed as a particular example of language coevolution. In our opinion, this conclusion opens new perspectives for further research in language competition, indicating that evolutionary approach can be very promising in the study of this phenomenon.

Acknowledgment

The author wishes to thank prof. Marek Skomorowski for his support and valuable suggestions.

References

Abrams, D. M. and Strogatz, S. H. (2003). Modelling the dynamics of language death, *Nature* **424**: 900.

Baronchelli, A., Felici, M., Loreto, V., Caglioti, E. and Steels, L. (2006). Sharp transition towards shared vocabularies in multi-agent systems, *Journal of Statistical Mechanics: Theory and Experiment* **2006**(06): P06014.

Baronchelli, A., Loreto, V. and Steels, L. (2008). In-depth analysis of the naming game dynamics: the homogeneous mixing case, *International Journal of Modern Physics C* **19**(05): 785–812.

Castelló, X., Eguíluz, V. M., Miguel, M. S., Loureiro-Porto, L., Toivonen, R., Saramäki, J. and Kaski, K. (2008). Modelling language competition: bilingualism and complex social networks, *in* A. D. M. Smith, K. Smith and R. F. i Cancho (Eds), *The Evolution of Language: Proceedings of the 7th International Conference (EVOLANG7)*, World Scientific, pp. 59–66.

Castelló, X., Loureiro-Porto, L. and Miguel, M. S. (2013). Agent-based models of language competition, *International Journal of the Sociology of Language* **2013(221)**: 21–51.

Gontier, N. (2017). What are the units of language evolution?, *Topoi* **8**(4): 1–19.

Hurford, J. R. (1989). Biological evolution of the saussurean sign as a component of the language acquisition device, *Lingua* **77**(2): 187–222.

Jansson, F., Parkvall, M. and Strimling, P. (2015). Modeling the evolution of creoles, *Language Dynamics and Change* **5**(1): 1–51.

Lipowska, D. (2011). Naming game and computational modelling of language evolution, *Computational Methods in Science and Technology* **17**(1): 41–51.

Lipowska, D. (2016). *Komputerowe modelowanie ewolucji języka (in Polish)*, 1 edn, Wydawnictwo Naukowe UAM, Poznań.

Moseley, C. (Ed.) (2010). *Atlas of the World's Languages in Danger*, 1 edn, UNESCO Publishing, Paris. Accessed on 29 Dec. 2018.
URL:`http://www.unesco.org/culture/en/endangeredlanguages/atlas`

Moulton, W. G. and Buccini, A. F. (2014). East germanic languages. Accessed on 29 Dec. 2018.
URL:`www.britannica.com/topic/East-Germanic-languages`

Mufwene, S. S. (2017). Creole languages. Accessed on 29 Dec. 2018.
URL:`www.britannica.com/topic/creole-languages`

Patriarca, M., Castelló, X., Uriarte, J. R., Eguíluz, V. M. and Miguel, M. S. (2012). Modeling two-language competition dynamics, *Advances in Complex Systems* **15**(03n04): 1250048.

Patriarca, M. and Leppännen, T. (2004). Modeling language competition, *Physica A* **338**: 296–299.

Schulze, C. and Stauffer, D. (2006). Computer simulation of language competition by physicists, *in* B. K. Chakrabarti, A. Chakraborti and A. Chatterjee (Eds), *Econophysics and Sociophysics: Trends and Perspectives*, John Wiley & Sons, Ltd, pp. 311–337.

Steels, L. (1995). A self-organizing spatial vocabulary, *Artificial Life* **2**(3): 319–332.

Steels, L. (1996). Emergent adaptive lexicons, *in* P. Maes, M. J. Mataric, J.-A. Meyer, J. Pollack and S. W. Wilson (Eds), *From Animals to Animals 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior*, The MIT Press, pp. 562–567.

Steels, L. and Loetzsch, M. (2012). The grounded naming game, *in* L. Steels (Ed.), *Experiments in Cultural Language Evolution*, John Benjamins Publishing Company, pp. 41–59.

Vogt, P. (2009). Modeling interactions between language evolution and demography, *Human Biology* **81**(2): 237–258.

Wang, W. S.-Y. and Minett, J. W. (2005). The invasion of language: emergence, change and death, *Trends in Ecology & Evolution* **20**(5): 263–269.

Wang, W. S.-Y. and Minett, J. W. (2008). Modeling endangered languages: The effects of bilingualism and social structure, *Lingua* **118**(1): 19–45.

Michał Grzejdziak (born 1997, Warsaw, Poland) is a bachelor of science in computer science and a student of computer mathematics at the Faculty of Mathematics and Computer Science at Jagiellonian University in Cracow. He is interested in applying computational methods in social sciences and humanities.