Naming Game and Computational Modelling of Language Evolution

Dorota Lipowska

Institute of Linguistics, Adam Mickiewicz University
60-371 Poznań, Poland
e-mail: lipowska@amu.edu.pl

Abstract: Computational modelling with multi-agent systems has become an important technique in studying language evolution. We present a brief introduction into this rapidly developing field, as well as our own contributions, which include an analysis of the evolutionary naming game model. In this model, communicating agents, which try to establish a common vocabulary, are equipped with an evolutionarily selected learning ability. Such a coupling of biological and linguistic ingredients results in an abrupt transition: upon a small change of the model control parameter, a poorly communicating group of agents with small learning abilities transforms into almost perfectly communicating group of agents with large learning abilities. Genetic imprinting of the learning abilities progresses through the Baldwin effect: initially linguistically unskilled agents learn a language, which creates a niche where there is an evolutionary pressure for the increase of learning ability. Under the assumption that communication intensity increases continuously with finite speed, the transition is split into several transition-like changes. It shows that the speed of cultural changes, that sets an additional characteristic time scale, might be yet another factor affecting the evolution of language. In our opinion, this model shows that linguistic and biological processes have a strong influence on each other and this influence certainly has contributed to an explosive development of our species.

Key words: computer modelling, agent systems, naming game, Baldwin effect, bio-linguistics

I. INTRODUCTION

I.1. Evolutionary Forces behind Language Development

The ability to use language distinguishes humans from all other species. Certain species have also developed some communication modes but of much smaller capabilities and complexity. For several decades various schools have tried to explain the emergence and development of language using diverse perspectives. Nativists argue that our language capacity is a collection of domain-specific cognitive skills that are somehow encoded in our genome. However, the idea of the existence of such a Language Acquisition Device or "language organ" (the term coined by their most prominent representative Noam Chomsky [1]), has been challenged by empiricists, who argue that linguistic performance of humans can be explained using domain-general learning techniques. Recently, Sampson [2] has even questioned the most appealing argument of nativists, which refers to the poverty of stimulus and apparently fast learning of grammar by children. An important issue of possible adaptative merits of language does not seem to be settled either. Non-adaptationists, again with Chomsky as the most famous representative [3], consider language as a side effect of other skills and thus claim that its evolution, at least at the beginning, was not attributable to any fitness advantage. A chief argument against the non-adaptationist stand is the observation that there is a number of costly adaptations that seem to support human linguistic abilities such as a large brain, a longer infancy period or descended larynx. Recently, in their influential paper Pinker and Bloom argued that, similarly to other complex adaptations, language evolution can be explained entirely by means of natural selection mechanisms [4]. Their paper triggered a number of works where language was examined from the perspective of evolutionary biology or game theory [5, 6]. In particular, Nowak et al. used some optimization arguments which might explain the origin of some linguistic universals [7]. They suggested that words appeared in order to increase the expressive capacity and

sentences (formed of words) limit memory requirements. The confrontation of nativists with empiricists and adaptationists with non-adaptationists does not seem to lead to consensus so far but certainly deepens our understanding of these problems [8].

Recently, a lot of works on language emergence have an evolutionary flavour. Such an approach puts some constraints on possible theories of the language origin. In particular, it rules out non-adaptationist theories, where language is a mere by-product of having a large and complex brain [9, 10]; see [4] for discussion. The emergence of language has also been listed as one of the major transitions in the evolution of life on Earth [11]. An interesting question is whether this transition was variation- or selection-limited [12]. In variation-limited transitions the required configuration of genes is highly unlikely and it takes a considerable amount of time for the nature to invent it. For selection-limited transitions the required configuration is easy to invent but there is no (or only very weak) evolutionary pressure that would favour it. Relatively large cognitive capacities of primates and their genetic proximity with humans suggest that some other species could also have been capable of developing language-like communication. Since they did not, it was perhaps due to a weak selective pressure. Such indirect arguments suggest that the emergence of language was selection-limited [12].

Some interesting results can be obtained by applying game-theory reasoning to one of the most basic problems of emerging linguistic communication, namely why do we talk and why do we exchange valuable and trustful information. Since speaking is costly (it takes time, energy and sometimes might expose a speaker to predators), and listening is not, such a situation seems to favour selfish individuals that would only listen but would not speak. Moreover, in the case of the conflict of interests the emerging communication system would be prone to misinformation or lying. The resolution of these dilemmas usually refers to the kin selection [13] or reciprocal altruism [14]. In other words, speakers remain honest because they are helping their relatives or they expect that others will do the same for them in the future. As an alternative explanation Dessalles [15] suggests that honest information is given freely because it is profitable – it is a way of competing for status within a group. Some related results on computational modelling of the honest cost-free communication are reported by Noble [16].

A necessary ingredient of language communication is learning. It is thus legitimate to ask whether Darwinian selection might account for the genetic hard-wiring of a Language Acquisition Device. Indeed, this (hypothetical)

organ is most likely responsible for some of the arbitrary (as opposed to the functional) linguistic structures. But for such an organ to be of any value, an individual has to acquire the language first. The inheritance of characteristics acquired during an individual lifetime is usually associated with long-discredited Lamarckian mechanism and thus considered to be doubtful. However, the relation between evolution and learning is more delicate and the attempts to clarify the mutual interactions of these two adaptive mechanisms have a long history. According to a purely Darwinian explanation, known as the Baldwin effect [17-19], there might appear a selective pressure in a population for the evolution of the instinctive behaviour that would replace the beneficial, but costly, learned behaviour [20]. Presumably, the Baldwin effect played an important role in the emergence and evolution of language but certain aspects of these processes still remain unresolved [21]. For example, a relatively stable environment is needed for the Baldwin effect to occur since rather slow evolutionary processes will otherwise not keep up with the fast changing environment. As the language formation processes are rather fast (in comparison to the evolutionary time scale), Christiansen and Chater questioned the role of adaptive evolutionary processes in the formation of arbitrary structures like Language Acquisition Device [22]. Actually, they suggest a much different scenario, where it is a language that adapted to human brain structures rather than vice versa.

I.2. Language as a Complex Adaptive System

From the above description it is clear that studying the emergence and evolution of language is a complex and multidisciplinary task and requires cooperation of not only linguists, neuroscientists, and anthropologists, but also experts in artificial intelligence, computer sciences or evolutionary biology [23]. One can distinguish two levels at which language can be studied and described [24] (Fig. 1). At the first level the description centers on the individual language users: their linguistic performance, language acquisition, speech errors, speech pathologies or brain functioning in relation with language processing. The language of each individual is slightly different. Nevertheless, within a certain population these individuals can efficiently communicate and that establishes the population level. At this level language is considered as an abstract system that exists in a sense separately from its individual users. There are numerous interactions between these two levels. Indeed, the linguistic behaviour of individuals depends on the language specific to the population they are

part of. And, as a feedback, the language used in a given population emerges from linguistic behaviour of individuals composing this population. Various processes shaping such a complex system operate at different time scales. The fastest dynamics operates at the individual level (onto-genetic time scale [25]), which includes, for example, language acquisition processes. Much slower processes, such as migrations of language populations, dialect formation or language extinctions, operate at the so-called glossogenetic time scale. The slowest processes are involved in the biological evolution of language users, which defines the phylogenetic time scale. Processes operating at these different time scales are not independent (Fig. 1). Biological evolution might change linguistic performance of individuals, which might affect the glossogenetic processes, which might in turn have an influence on phylogenetic development of individuals. For example, a mutation that changes the vocal ability of a certain individual, if spread in his/her population, might lead to a dialect formation or a language extinction. Such population-level processes might change the selective pressure that individual language users are exposed to, which might affect phylogenetic processes, thus closing the interaction loop.

Various levels of descriptions and processes operating at several time scales suggest that undoubtedly, to describe
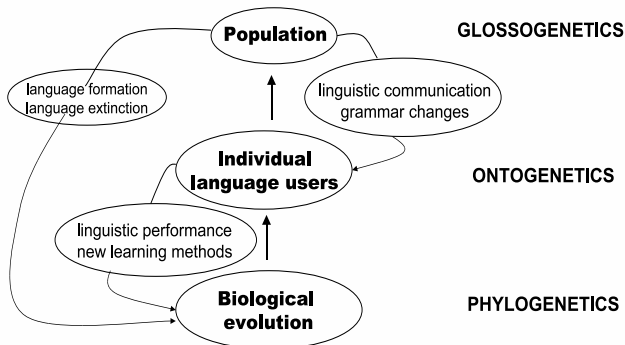

Fig. 1. Language as a complex adaptive system. Many different processes involved in language evolution operate at various levels and time scales, but nevertheless strongly influence each other. Individual-level processes (ontogenetics), such as, e.g., language acquisition or acts of communication, i.e. those determined mainly by interactions between individual language users, are relatively fast. Much slower are population-level processes (glossogenetics), such as language formations or extinctions, grammar changes, and migrations. To obtain a complete description one has also to consider biological evolution of language users (phylogenetics), the processes of which (the slowest ones among those affecting language evolution) determine individuals via their genetic endowment. Various processes at individual and population levels affect the fitness landscape, which influences the biological evolution level. Similarly, individual language users are affected by populational and evolutionary processes

adequately the language evolution, extremely complex models must be used. Correspondingly, analyzing such models and predicting their behaviour must also be very difficult. It is known that some phenomena involving feedback interactions might be described in terms of nonlinear differential equations, such as, for example, Lotka-Volterra equations describing interacting populations. The behaviour of such nonlinear equations is often difficult to predict, since abrupt changes even of the qualitative nature of solutions might take place. Language evolution, however, seems to be a much more complex issue than ecological problems of interacting populations, and its description in terms of differential equations would be much more complicated if at all feasible.

Recently, the most promising and frequently used approach to examine such systems is computational modelling of multi-agent systems. Using this method one examines a language that emerges in a bottom-up fashion as a result of interactions within a group of the so-called agents equipped with some linguistic functions. Then one considers language as a complex adaptive system that evolves and complexifies according to biologically inspired principles such as selection and self-organization [26]. Thus, the emerging language is not static but evolves in a way that hopefully is similar to human language evolution. Of course, using such an approach one cannot explain all intricacies of human languages. A more modest goal would be to understand some rather basic features that are common to all languages such as meaning-form mappings, origin of linguistic coherence (among agents without central control and global view), or coevolutionary origin of grammar and meaning.

Within the multi-agent approach, two groups of models can be distinguished. In the first one, originating from the so-called iterated learning model, one is mainly concerned with the transmission of language between successive generations of agents [27, 28]. Agents classified as teachers produce some expressions, which are passed to learners. Using statistical learning techniques such as, e.g., neural networks, learners try to infer a grammar producing such expressions. Next, the teachers are replaced by the learners while a new population of learners is introduced, and the procedure is iterated. The most important issue that the iterated learning model has successfully addressed is the transition from a holistic language (in which complex meanings are expressed in unstructured arbitrary forms) to a compositional language (in which complex meanings are expressed in structured forms). However, since such procedure is computationally relatively demanding and the number of communicating agents is thus typically very small (for example, one teacher and one learner), the prob-

lem of the emergence of linguistic coherence must be neglected in this approach. To tackle this problem Steels introduced a naming game model [29]. In this approach one examines a population of agents trying to establish a common vocabulary for a certain number of objects present in their environment. The change of generations is not required here since the emergence of a common vocabulary is a consequence of the communication processes between agents, and they are not divided into teachers and learners but take the roles of a speaker or a hearer alternately.

It seems that the iterated learning model and the naming game model are at two extremes: the first model emphasizes the generational turnover while the second one concentrates on one-generation (cultural) interactions. Since in language evolution both aspects are present, it is desirable to examine models that combine evolutionary and cultural processes. Recently we have introduced such a model [30] and one of the objectives of the present paper is to provide further analysis of its behaviour based on more extensive simulations. Our model captures all three basic aspects of language: learning, culture, and evolution. In our model agents try to establish a common vocabulary like in the naming game model, but in addition they can breed, mutate, and die. Moreover, they are equipped with an evolutionary trait: learning ability. When communication between agents is sufficiently frequent, cultural processes create a niche in which a larger learning ability becomes advantageous. Thus learning abilities increase, and in turn their large values make the cultural processes more efficient. As a result the model undergoes an abrupt biolinguistic transition where both linguistic performance and abilities of agents change very rapidly [30]. One of the main results reported in this paper is that under the plausible assumption of the intensity of communication increasing continuously in time, this bio-linguistic transition is replaced with a series of fast, transition-like changes. In our opinion, the proposed model shows that linguistic and biological processes have a strong influence on each other, which has certainly contributed to an explosive development of our species. That learning in our model modifies the fitness landscape of a given agent and facilitates the genetic accommodation of learning ability is actually a manifestation of the much debated Baldwin effect.

II. MODEL

In our model we consider a set of agents located at sites of a square lattice of a linear size L. The agents are trying to establish a shared vocabulary referring to a single object present in their environment. An assumption that agents communicate only about a single object does not seem to restrict the generality of our considerations and has already been used in some other studies of naming game [31, 32] or language change [33] models. A randomly selected agent takes the role of a speaker that communicates a word chosen from its inventory to a hearer that is randomly selected among nearest neighbours of the speaker. The hearer tries to recognize the communicated word, namely it checks whether it has the word in its inventory. A positive or negative result translates into communicative success or failure, respectively. In some versions of the naming game model [31, 32], a success means that both agents retain in their inventories only the chosen word, while in the case of failure the hearer adds the communicated word to its inventory.

To implement the learning ability, we have modified this rule and assigned weights w_i ($w_i > 0$) to each i-th word in the inventory. The speaker selects then the i-th word with the probability $w_i/\Sigma_j w_j$ where summation is over all words in its inventory (if its inventory is empty, it creates a word randomly). If the hearer has the word in its inventory, it is recognized. In addition, each agent k is characterized by its learning ability l_k ($0 < l_k < 1$), which is used to modify weights. In the case of success both speaker and hearer increase the weights of the communicated word by their respective learning abilities. In the case of failure the speaker subtracts its learning ability from the weight of the communicated word. If after the subtraction the weight becomes negative, the corresponding word is removed from the speaker's inventory. The hearer in the case of failure, i.e., when it does not have the word in its inventory, adds the communicated word to the inventory with a unit weight.

Apart from communication, agents in our model evolve according to the population dynamics: they can breed, mutate, and eventually die. To specify the intensity of these processes, we have introduced a communication probability p. With the probability p the chosen agent becomes a speaker, and with the probability $1 - p$ it undergoes appropriate vital processes. Namely, the agent either dies with the probability $1 - p_{\text{surv}}$, or survives with the probability p_{surv}, where $p_{\text{surv}} = \exp(-at)[1 - \exp(-b \Sigma_j w_j/\langle w \rangle)]$, and $a = 0.05$ and $b = 5$ are certain parameters whose role is to ensure a suitable speed of population turnover. Moreover, t is the age of an agent and $\langle w \rangle$ is the average (over agents) sum of weights. The above formula takes into account both the agent's linguistic performance (the bigger $\Sigma_j w_j$ the larger p_{surv}) and its age (the older it is the smaller its p_{surv} is). If the agent survives, it breeds, provided that there is an

empty site in its nearest neighbourhood. The offspring typically, i.e. with the probability $1 - p_{mut}$, inherits the parent's learning ability and the word of the highest weight from its inventory (in the offspring's inventory the initial weight of the word equals one). With a small probability p_{mut} a mutation takes place and the learning ability of an off-spring is randomly selected. With the same probability p_{mut} an independent check is made whether to mutate the inherited word. A diagram illustrating the dynamics of our model is given in the Appendix. Let us also notice that the behaviour of our model which is described below is to some extent robust with respect to some modifications of its rules. For example, qualitatively the same behaviour is observed for modified values of the parameters a and b, a different form of the survival probability p_{surv} (provided it is a decreasing function of the agent's age t and an increasing function of the quality of its linguistic performance $\Sigma_j w_j$), or different breeding and/or mutation rules.

III. RESULTS

To examine the properties of the model, numerical simulations have been performed. Most of the results have been obtained for $L = 60$ and $p_{mut} = 0.001$, but simulations for $L = 80$ or $p_{mut} = 0.01$ have led to similar results. A typical simulation starts on a lattice with each site occupied by an agent having in its inventory a single word chosen randomly for each agent and assigned a unit weight. Unless specified otherwise, the learning ability of each agent in the initial configuration is also chosen randomly.

III.1. Bio-linguistic Transition

An important parameter of the model is the communication probability p that specifies the intensity of communication attempts in comparison with population changes. In general, for small p the model remains in a phase of linguistic disorder with only small clusters of agents using the same language. We define the language of an agent as the highest-weight word in its inventory. Such a definition means that agents using the same language usually (but not always) use a mutually recognizable word, which ensures a relatively large rate of communication successes for such agents. A typical distribution of languages in this disordered small-p phase is shown in the upper left panel of Fig. 2, where agents using the same language are drawn with the same shade of grey. Upon increasing the communication probability p, the clusters of

agents only slightly increase, but after reaching a certain threshold an abrupt transition takes place and the model enters a phase of linguistic coherence with almost all agents belonging to the same cluster (Fig. 2, lower left panel).
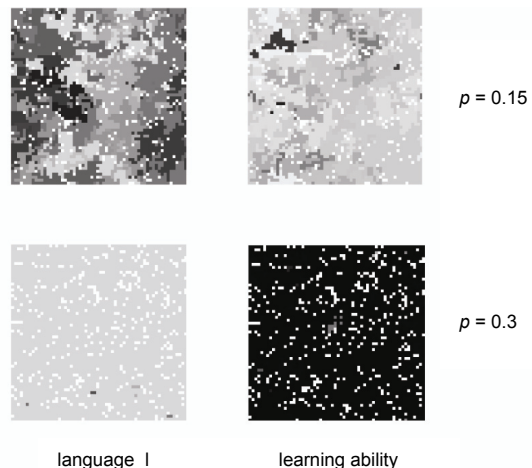

$p = 0.15$

$p = 0.3$

language l learning ability

Fig. 2. Exemplary configurations of the evolutionary naming game model with $L = 60$ and $p_{mut} = 0.001$. Agents using the same language (left panels) or having the same learning ability (right panels) are marked with the same shade of grey (white: $l = 0$, black: $l = 1$, i.e., the darker – the greater learning ability). In the small-p phase (upper panels) communication acts are infrequent, and agents using the same language (left) or having the same learning ability (right) form only small clusters. In this phase the communication success rate s and the average learning ability l are small (see also Figs. 3-4). In the large-p phase (lower panels) more frequent acts of communication result in the emergence of a common language. Thus, almost all agents use the same language and, moreover, have the same (and great) learning ability

To examine the nature of this transition, we have measured the communication success rate s defined as an average over agents and simulation time of the fraction of successes with respect to all communication attempts (Fig. 3). Moreover, we have measured the average learning ability l (Fig. 4). Since death, birth, and mutation processes depend on the linguistic performance of agents, during the evolution of the model some agents will be more favoured than others. As a result, the average learning ability l and the success rate s turn out to depend on the communication probability p. One can notice that upon increasing p the abrupt transition takes place around $p = 0.23$, where both the communication success rate s and the learning ability l jump rapidly. Moreover, upon decreasing p this transition takes place at a much lower value, namely around $p = 0.15$. Such a hysteretic behaviour indicates that the transition in our model is discontinuous. Let us also emphasize, that although s and l seem to exhibit a similar p-dependence,

they are of much different nature – while the success rate characterizes the overall linguistic performance of a given agent (during its lifetime), the learning ability is an inherited property, which does not change during the lifetime of an agent.
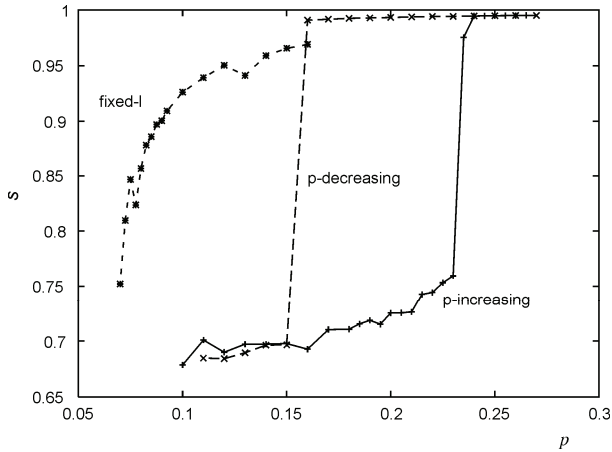

Fig. 3. The communication success rate s as a function of the communication probability p. Calculations were made for the system size $L = 60$ and the mutation probability $p_{mut} = 0.001$. Simulation time for each value of p was typically equal to 10^5 steps with 3×10^4 steps discarded for relaxation. A step is defined as a single, on average, update of each site. For simulations with decreasing p, we first relaxed the system until a mono-language state was reached (with s and l close to unity). In the left part of the graph the data are from simulations with fixed $l = 0.5$
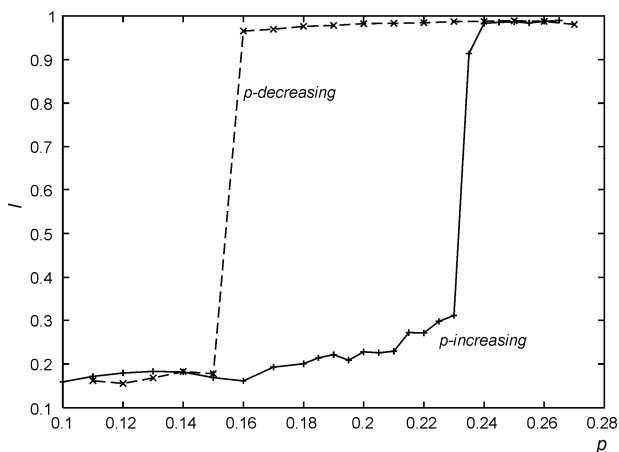

Fig. 4. The average learning ability l as a function of the communication probability p. Details of the simulations are the same as in Fig. 3

We have also examined the behaviour of the model with the learning ability kept fixed during entire simulations. In this case there is also a phase transition between disordered and linguistically coherent phases but this time

the transition is much smoother and there is no indication of the hysteretic behaviour (Fig. 3). To get further insight into the behaviour of our model, we have measured the fraction f_m of agents using the language with the largest number of users. Simulations show that for the learning ability kept fixed f_m decreases also in a much smoother way (Fig. 5). Moreover, its variance has a pronounced peak at the transition point, which this time takes place around $p = 0.07$ (Fig. 6). Such large fluctuations of f_m (and the variance of s shows a similar behaviour) in the vicinity of the transition point and an absence of a jump suggest that it might be a continuous transition.
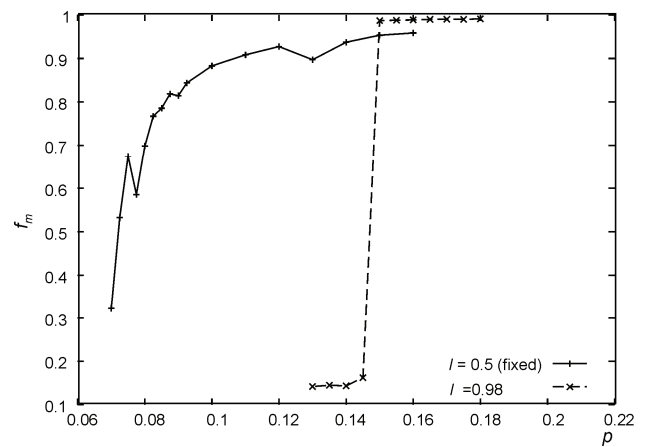

Fig. 5. The fraction f_m of agents using the language with the largest number of users as a function of the communication probability p. For simulations with the learning ability not kept fixed, we started from the configuration with all agents having the same word in their repositories and the learning ability set to 0.98. Such a choice of the initial state leads to only minor differences in comparison with simulations in Figs. 3-4
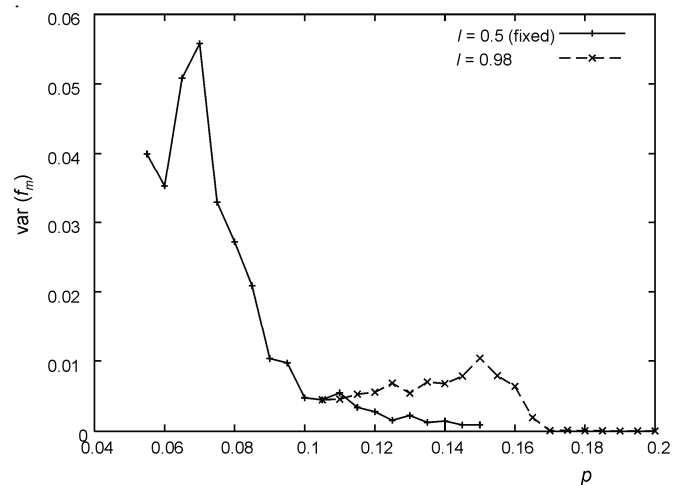

Fig. 6. The variance of f_m. Details of simulations are the same as in Fig. 5
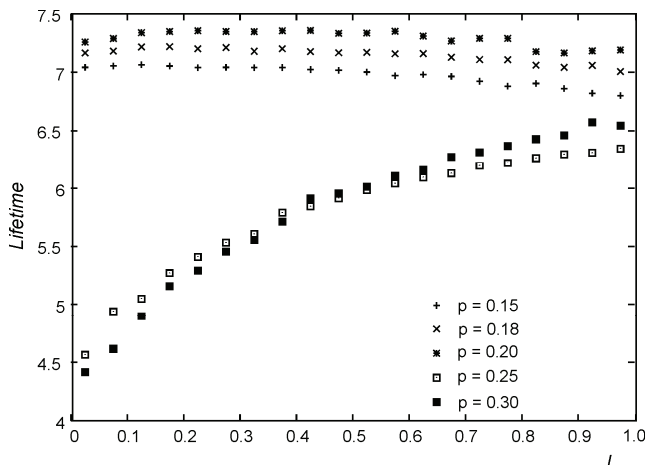
Fig. 7. Lifetime of agents as a function of the learning ability *l* for several values of the communication probability *p*. One can notice that in a predictable environment (large-*p* phase) having a great learning ability is advantageous. In a random environment (low-*p* phase), the lifetime of an agent is almost independent on its learning ability

A noticeable difference between small-p and large-*p* phases appears in the learning-ability dependence of lifetime of agents (Fig. 7). One can see that in the large-*p* phase, where neighbouring agents are likely to use the same language, having a large learning ability increases the agent's lifetime (faster learning enables faster accommodation to the predictable linguistic environment). On the other hand, in the small-*p* phase (i.e., in the random linguistic environment) the lifetime is almost independent on the learning ability. Before presenting computational results concerning the dynamics of our model, let us notice that sudden transitions in linguistic models were also reported in some other models [7, 38].

III.2. Dynamic Behaviour

Because each agent is characterized by its learning ability, states which are homogeneous (namely such that a majority of agents use the same language) but of different average learning abilities are not equivalent. As a result, evolution of the model in an intricate way depends on the initial configuration and the parameters p and p_{mut}. This is particularly transparent in the range $0.15 < p < 0.25$, where the model exhibits a hysteretic behaviour (for $p_{mut} = 0.001$). An example that shows the dependence of the behaviour of the model on its initial configuration is shown in Fig. 8. In a center of a lattice of size $L = 60$, we have created a square seed of 100 agents having the same learning ability (0.98) and the same word in their inventories. This seed is surrounded by $60 \times 60 - 100 = 3500$ agents of a smaller

learning ability (0.5). As can be seen in Fig. 8, the evolution depends on whether the surrounding agents use the same language as those in the seed (homogenous case) or whether their repositories contain initially random words
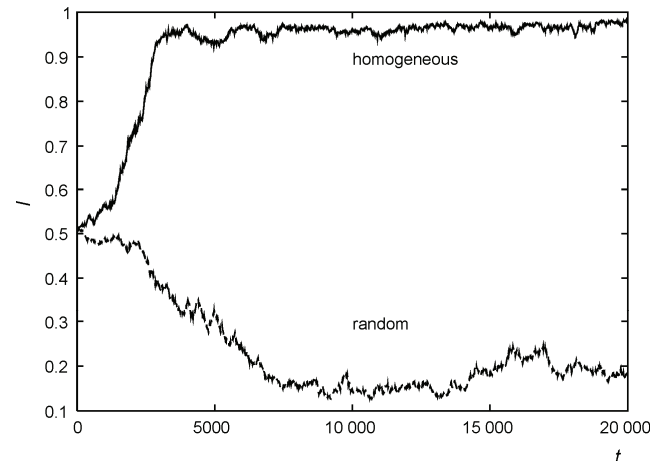

Fig. 8. Time evolution of the learning ability *l* for the $L = 60$ model with the seed of 100 agents with $l = 0.98$ surrounded by 3500 agents with $l = 0.5$. The course of the evolution depends on the initial state (here: inventories) of surrounding agents (see main text for the detailed description)

(random case). In the first case the system ends up in a homogenous state where a majority of agents use the same language and have the same learning ability (0.98). In the second case the model evolves toward a multi-language state with much smaller learning abilities. In such a setup the size of the seed or the learning abilities of the surrounding agents are also important parameters that might affect the course of the evolution of the model. For example, we observed that also in the homogenous case but with surrounding agents having an even smaller learning ability (0.3), the model evolved toward a multi-language state.

III.3. The Baldwin Effect

The fact that both the communication success rate *s* and the average learning ability *l* have a jump at the same value of *p* (Figs. 3-4) shows that communicative and biological ingredients strongly influence each other, which leads to a single and abrupt transition. In our model successful communication requires learning. A new-born agent communicating with some mature agents who already worked out a certain (common in this group) language will increase the weight of a corresponding word. As a result, in

subsequent acts of communication the agent will use mainly this word. In what way might such learning get coupled with evolutionary traits? The explanation of this phenomenon is known as the Baldwin effect. Although at first sight it looks like a discredited Lamarckian phenomenon, the Baldwin effect is actually purely Darwinian [34, 21]. There are usually some benefits related with the task a given species has to learn and there is a cost of learning this task. One can argue that in such case there is some kind of an evolutionary pressure that favours individuals for which the benefit is larger or the cost is smaller. The evolution will then lead to the formation of species where the learned behaviour becomes an innate ability. It should be emphasized that the acquired characteristics are not inherited. What is inherited is the ability to acquire the characteristics (the ability to learn) [20]. In the context of the language evolution the importance of the Baldwin effect was suggested by Pinker and Bloom [4]. Perhaps this effect is also at least partially responsible for the formation of the Language Acquisition Device (and in our model the learning ability l, as a feature heritable and unchangeable during a phenotypic development, might be regarded as approximately corresponding to the LAD). However, many details concerning the role of the Baldwin effect in the evolution of language remain unclear [35].

We have already argued [30] that in our model the Baldwin effect is also at work. Let us consider a population of agents when the communication probability p is below the threshold value ($p < p_c \approx 0.23$). In such a case the learning ability remains at a rather low level (since clusters of agents using the same language are small, it does not pay off to be good at learning the language of neighbours). Now, let us increase the value of p above the threshold value. More frequent communication changes the behaviour dramatically. Apparently, clusters of agents using the same language are now sufficiently large and it pays off to have a large learning ability because that increases the success rate and thus the survival probability p_{surv}. Let us notice that p_{surv} of an agent depends on its linguistic performance ($\Sigma_j w_j$) rather than its learning ability. Thus clusters of agents of good linguistic performance (learned behaviour) can be considered as niches that direct the evolution by favouring agents with large learning abilities, which is precisely the Baldwin effect. It should be noticed that linguistic interactions between agents (whose rate is set by the probability p) are typically much faster than evolutionary changes (set by p_{mut}) and such an effect has also been observed in simulations [30].

As a result of a positive feedback (large learning ability enhances communication, which enlarges clusters, which even more favours an increased learning ability) a discon-

tinuous transition takes place both with respect to the communication success rate and the average learning ability (Figs. 3-4). An interesting question is whether such a behaviour is of any relevance in the context of human evolution. It is obvious that development of language, which probably took place somewhere around 10^5 years ago, was accompanied by important anatomical changes such as fixation of the so-called speech gene (FOXP2), descended larynx or enlargement of brain [36]. Linguistic and other cultural interactions that were already emerging in early hominid populations were certainly shaping the fitness landscape and that could direct the evolution of our ancestors via the Baldwin effect.

Since it is plausible that communication attempts in the human history were gradually becoming more frequent (and important), it is natural to simulate our model with the communication probability p increasing continuously in time. In the initial population, the agents' learning abilities should be rather small. Since human populations are quite homogeneous with respect to linguistic abilities, it is desirable that the model dynamics should lead to an l-homogeneous state, i.e., a state where the majority of agents have the same, and large, learning abilities. Although many languages are now at the verge of extinction and one cannot exclude that in the future humans will use only one language, at least at present many languages exist. Thus in the model's final state, or at least in a transient, although long-lasting state, there should be many languages (or rather several, taking into account limitations of the simulations). The results of our simulations can be seen in Fig. 9. Initially, the learning abilities have been set as random numbers distributed uniformly from the interval (0, 0.1). One can notice that around $t = 5 \times 10^4$ agents with the learning ability close to 0.1 dominate the system (the fraction of agents having the most frequently occurring learning ability $f_l \approx 1$). However, with time and increasing communication probability p, such a small learning ability becomes insufficient, and around $t = 12 \times 10^4$ agents with learning ability close to 0.3 dominate the system. However, f_m, i.e., the fraction of agents whose language is used by the largest number of agents, still remains close to 0, which means that even the most widespread language is used only by a few agents. Around $t = 15 \times 10^4$ the next transition takes place and agents with large learning abilities dominate the system. Around that time f_m starts to increase, which means that some languages diffuse while some become extinct. Since almost all agents have the same language ability, all languages are dynamically equivalent. Eventually, the system reaches the state where almost all agents use the same language (f_m, $s \approx 1$), however, the time needed to reach such a state might be quite long.

In Fig. 9 the behaviour of the model in the interval $15 \times 10^4 < t < 2 \times 10^5$ resembles the current stage of the human language evolution: a maximum learning ability dominating the population and several (not too many and not too few) languages in use. Before this final state, some plateaus can be distinguished separated by relatively rapid transitions. Such a behaviour differs from the single-step scenario seen in the simulations where p increases in finite steps but is kept constant during measurements (Figs. 3-4).
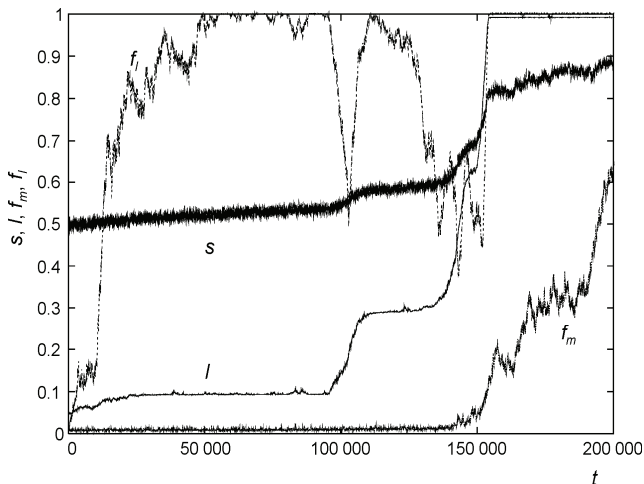

Fig. 9. Time evolution of the model characteristics upon the linear in time increase of the communication probability p from 0.1 to 0.5 (the lattice size $L = 60$). We measured the success rate s, the average learning ability l, the fraction of agents whose language is used by the largest number of agents f_m, and the fraction of agents having the most frequently occurring learning ability f_l. One can see that around $t = 15 \times 10^4$ both f_l and l become close to unity, which means that almost every agent has the same maximum learning ability. Further evolution gradually eliminates less widespread languages and leads to the state where almost all agents use the same language ($f_m, s \approx 1$)

Presumably the multi-step behaviour is a consequence of a continuous increase of p. As already discussed (Sec. I.2), basic factors that determine the evolution of language set some characteristic time scales of the corresponding processes. Thus, individual learning would take dozens of years, culture hundreds of years, and biological evolution most likely dozens of thousands of years. The speed of increase of p that might be interpreted as a speed of cultural changes has yet another characteristic time scale and our work shows that this scale might influence the evolution of language. Certainly, further research is needed to examine in more detail an intricate influence of learning, culture, and biological evolution on language.

IV. CONCLUSIONS

In the present paper we have examined an evolutionary naming game model. Simulations show that the coupling of linguistic and evolutionary ingredients produces a discontinuous transition and that learning can direct the evolution towards greater linguistic abilities (the Baldwin effect). However, under perhaps slightly more realistic assumptions, i.e. when the communication probability increases continuously, this transition is split into a series of transitions. This result shows that the speed of cultural changes might be yet another factor affecting the evolution of language and setting an additional characteristic time scale. The present model is not very demanding computationally. It seems to be possible to consider agents talking about more than one object, or to examine statistical properties of simulated languages such as, for example, distributions of their lifetimes or of the number of users. One can also study effects like diffusion of languages, the role of geographical barriers, or formation of language families. There is already an extensive literature documenting linguistic data as well as various computational methods of modelling, for example, a competition between natural languages [37-39]. The dynamics of the present model, based on an elementary act of communication, perhaps offers a more natural description of dynamics of languages than some other approaches that often use some kind of coarse-grained dynamics.
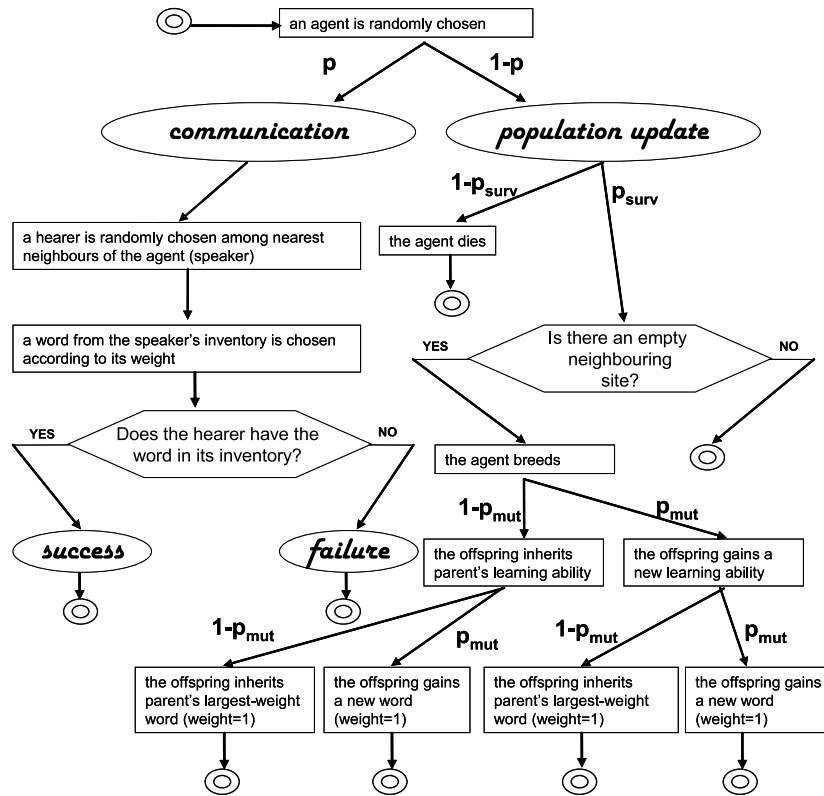
References

[1] N. Chomsky, *Aspects of the Theory of Syntax*. MIT Press, Cambridge 1965.

[2] G. Sampson, *Educating Eve: The 'Language Instinct' Debate*. Cassell, London 1997.

[3] N. Chomsky, *Language and Mind, Harcourt Brace Jovanovich*. San Diego 1972.

[4] S. Pinker, P. Bloom, *Natural language and natural selection*. Behavioral and Brain Sciences 13 (4), 707-784 (1990).

[5] R.S. Jackendoff, *Languages of the Mind*. MIT Press, Cambridge 1992.

[6] C. Knight et al. (eds.), *The Evolutionary Emergence of Language: Social Function and the Origin of Linguistic Form*. Cambridge University Press, Cambridge 2000.

[7] M.A. Nowak, N.L. Komarova, *Towards an evolutionary theory of language*. Trends in Cognitive Sciences 5 (7), 288-295 (2001). M.A. Nowak, D.C. Krakauer, *The evolution of language*. Proc. Natl. Acad. Sci. USA 96 (14), 8028-8033 (1999).

[8] K. Smith, *The Transmission of Language: models of biological and cultural evolution*. Ph.D. thesis, The University of Edinburgh, Edinburgh 2003.

[9] S.J. Gould, *The limits of adaptation: Is language a spandrel of the human brain?* Paper presented to the Cognitive Science Seminar, Center for Cognitive Science, MIT 1987.

[10] S.J. Gould, R.C. Lewontin, *The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme.* Proc. R. Soc. Lond. B 205 (1161), 581-598 (1979).

[11] J. Maynard Smith, E. Szathmáry, *Major Transitions in Evolution.* Freeman, Oxford 1995.

[12] S. Számadó, E. Szathmáry, *Selective scenarios for the emergence of natural language.* Trends in Ecology & Evolution 21(10), 555-561 (2006).

[13] W.D. Hamilton, *The genetical evolution of social behaviour I and II.* Journal of Theoretical Biology 7, 1-16, 17-52 (1964).

[14] R.L. Trivers, *The evolution of reciprocal altruism.* Quarterly Review of Biology 46, 35-57 (1971).

[15] J.-L. Dessalles, *Altruism, status, and the origin of relevance.* In: J.R. Hurford et al. (eds.), *Approaches to the Evolution of Language: Social and Cognitive Bases.* Cambridge University Press, Cambridge, 130-147 (1998).

[16] J. Noble, *Cooperation, competition and the evolution of prelinguistic communication.* In: Knight et al. (eds.) *Evolutionary Emergence of Language.* Cambridge University Press, Cambridge, 40-61 (2000).

[17] J.M. Baldwin, *A new factor in evolution.* American Naturalist 30, 441-451, 536-553 (1896).

[18] G.G. Simpson, *The Baldwin effect.* Evolution 7, 110-117 (1953).

[19] B.H. Weber, D.J. Depew (eds.), *Evolution and Learning – The Baldwin Effect Reconsidered.* MIT Press, Cambridge 2003.

[20] P.D. Turney, *Myths and legends of the Baldwin effect.* In: T. Fogarty, G. Venturini (eds.) *Proceedings of the ICML-96* (13th International Conference on Machine Learning, Bari, Italy), 135-142 (1996).

[21] H. Yamauchi, *Baldwinian Accounts of Language Evolution.* Ph.D. thesis, The University of Edinburgh, Edinburgh 2004.

[22] M.H. Christiansen, N. Chater, *Language as shaped by the brain.* Behavioral and Brain Sciences 31, 489-558 (2008).

[23] M.A. Nowak, N.L. Komarova, P. Niyogi, *Computational and evolutionary aspects of language.* Nature 417, 611-617 (2002).

[24] B. de Boer, *Computer modelling as a tool for understanding language evolution.* In: Gonthier et al. (eds.) *Evolutionary Epistemology, Language and Culture – A non-adaptationist, systems theoretical approach.* Springer, Dordrecht, 381-406 (2006).

[25] S. Kirby, *Natural language from Artificial Life.* Artificial Life 8 (2), 185-215 (2002).

[26] L. Steels, *Iterated Learning versus Language Games. Two models for cultural language evolution.* In: C. Hemelrijk, E. Bonabeau (eds.) *Proceedings of the International Workshop of the Self-Organization and Evolution of Social Behaviour.* University of Zurich, Switzerland, Zurich 2002.

[27] S. Kirby, J. Hurford, The emergence of linguistic structure; *An overview of the Iterated Learning Model.* In: A. Canelosi, D. Parisi (eds.) *Simulating the Evolution of Language.* Springer-Verlag, Berlin p. 121-148, 2001.

[28] H. Brighton, *Compositional syntax from cultural transmission.* Artificial Life 8 (1), 25-54 (2002).

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

[30] A. Lipowski, D. Lipowska, *Bio-linguistic transition and the Baldwin effect in the evolutionary naming game model.* International Journal of Modern Physics C, 19 (3), 399-407 (2008).

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

[32] L. Dall'Asta, A. Baronchelli, A. Barrat, V. Loreto, *Nonequilibrium dynamics of language games on complex networks.* Physical Review E 74, 036105 (2006).

[33] D. Nettle, *Using social impact theory to simulate language change.* Lingua 108, 95-117 (1999). D. Nettle, *Is the rate of linguistic change constant?* Lingua 108, 119-136 (1999).

[34] G. Hinton, S. Nowlan, *How learning can guide evolution.* Complex Systems 1, 495-502 (1987).

[35] S. Munroe, A. Cangelosi, *Learning and the evolution of language: The role of cultural variation and learning costs in the Baldwin effect.* Artificial Life 8 (4), 311-339 (2002).

[36] C. Holden, *The origin of speech.* Science 303, 1316-1319 (2004).

[37] D. Abrams, S.H. Strogatz, *Modelling the dynamics of language death.* Nature 424, 900 (2003).

[38] C. Schulze, D. Stauffer, S. Wichmann, *Birth, survival and death of languages by Monte Carlo simulation.* Communications in Computational Physics 3 (2), 271-294 (2008).

[39] P.M.C. de Oliveira, D. Stauffer, S. Wichmann, S.M. de Oliveira, *A computer simulation of language families.* Journal of Linguistics 44 (3), 659-675 (2008).

APPENDIX

The following block diagram illustrates the rules of the evolutionary naming game model. Updating inventories in the case of success or failure is described in the text.

DOROTA LIPOWSKA received her M.Sc. in Mathematics (1981) and her Ph.D. in Arts (Linguistics; 1988), both from Adam Mickiewicz University (Poznań, Poland). Currently she is a lecturer at the same university, Institute of Linguistics (Department of Applied Logic). Her research interests concern computational linguistics as well as modelling the emergence and the evolution of language.