

Evolution of language diversity: Why fitness counts

Zach Solan, Eytan Ruppin, David Horn

Sackler Faculty of Exact Sciences

Tel Aviv University

Tel Aviv, Israel 69978

{*rsolan,ruppin,horn*}@post.tau.ac.il

Shimon Edelman

Department of Psychology

Cornell University

Ithaca, NY 14853, USA

se37@cornell.edu

July 4, 2003

Abstract

We examined the role of fitness, commonly assumed without proof to be conferred by the mastery of language, in shaping the dynamics of language evolution. To that end, we introduced island migration (a concept borrowed from population genetics) into the shared lexicon model of communication (Hurford, 1989; Nowak et al., 1999). The effect of fitness in language coherence was compared to a control condition of neutral drift. We found that in the neutral condition (no coherence-dependent fitness) even a small migration rate – less than 1% – suffices for one language to become dominant, albeit after a long time. In comparison, when fitness-based selection is introduced, the subpopulations stabilize quite rapidly to form several distinct languages. Our findings support the notion that language confers increased fitness. The possibility that a shared language evolved as a result of neutral drift appears less likely, unless migration rates over evolutionary times were extremely small.

1 Introduction

Many of the recent studies that explored the evolution of language in a population of communicating agents assume that the possession of a common language increases the fitness of its users, e.g., (Cangelosi, 2001; Reggia et al., 2001; Nowak et al., 1999; Smith, 2001). Language-conferred fitness is defined as the increase in individual survival probability that stems from successful communication: a group whose members communicate well is supposed to leave, on the average, more offspring. Although the assumption that language increases human evolutionary fitness is intuitively appealing, its validity needs to be examined in the light of the indications that a coherent language can emerge in the absence of language-related fitness (Kirby, 2000; Olliphant, 1999; Olliphant and Batali, 1997; Briscoe, 2003), e.g., through a combination of genetic drift and random population diffusion.

The present paper examines and compares the effects of fitness and of neutral drift on the emergence of coherent languages. Our approach is motivated both by communication models employed in studies of language evolution (Nowak et al., 2000; Nowak et al., 1999; Nowak and Krakauer, 1999; Komarova and Nowak, 2001; Kirby, 2000; Kirby, 2002; Batali, 1998), and by migration models that are widely used in population genetics, such as the island model (Wright, 1931) and migration matrix (Bodmer and Cavalli-Sforza, 1968). Specifically, we integrate communication and migration effects by constructing a system of several distinct populations that exchange migrants at a controlled rate; within each population, evolutionary dynamics of language is governed by the communication model.

The spread of languages within and between populations in this model can be compared to empirical data on linguistic diversity, a field of study that has been drawing increasing attention over the recent years (Nichols, 1992; Nettle, 1999). Although there are as many as 6500 known languages, about 95% of the world population speaks only 100 languages. The distribution of languages around the world is far from uniform (Grimes, 2001). Hundreds of languages have evolved in Africa and Oceania, while very few evolved in North America and Europe. Examining the distribution of languages, most of which are nearly extinct, can shed light on the evolutionary forces that were involved in their emergence.

The markedly uneven geographic distribution both of the emergence of languages and of their disappearance suggests that powerful dynamic mechanisms are at work. The evolutionary dynamics behind linguistic diversity can be explored using computational simulations, as in (Nettle, 1999) and (Livingstone, 2000). Nettle's work studied vowel learning in the presence of migration, in a population distributed over a 7×7 spatial grid, with 20 individuals at each position. Each new offspring learned two continuous parameters (the first two formants), by sampling their values among all the adults in its group. Individuals thus learned the approximate average values of the parameters in their group. Each simulation typically lasted for several hundreds of generations. To simplify the calculations and to keep the group sizes constant, migration always involved an exchange of individuals between two groups.

Nettle examined the dynamics of this system under three modes of fitness influence. The first of those was the neutral mode, corresponding to an absence of selection pressure. In that mode the individuals simply learned by averaging the values of the two continuous parameters among the adults in their social group. The second mode was functional selection, based on the notion that some variables may confer fitness because they make language easier to learn or use. In Nettle's version of the functional selection mode, if the distance between any two phonemes is less than a critical value then one of the phonemes is moved randomly in the formant space until the critical distance is reestablished (note that this process amplifies within-group differences). The third mode of fitness influence corresponded to raising the "social status" of some individuals in each group, and having the new offspring learn selectively from the high-ranking adults in their group. Nettle found that in the absence of social selection even a very low level of migration destroys local diversity. Social selection was found to contribute very strongly both to diversity and to its

stability in the face of higher migration rates. Finally, functional selection was seen to amplify and reinforce diversity, but could not bring it about on its own.

Livingstone's work (Livingstone, 2000; Livingstone and Fyfe, 2000) suggests that no functional benefits are required to create linguistic diversity and that language diversity may be a natural consequence of the adaptively neutral cultural evolution of language. Livingstone studied a computational model based on the evolutionary dynamics of a population of communicating individuals. These were simulated by simple artificial neural networks that mapped signals to meanings (once an agent is presented with a meaning, a signal is produced, and vice-versa). In contrast to Nettle's work, Livingstone's study did not address the issue of migration forces and was based solely on the transmission of signal-meaning maps between generations.

2 The model

In the present work, we extend Nettle's migration matrix model (controlling the population exchange between spatially separated "islands") by introducing a measure of the individual fitness based on communication success (Hurford, 1989; Nowak et al., 1999). Thus, in our model both the individual and the social fitness are determined by the communication performance, rather than by external parameters. In addition, we allow true matrix-controlled migration (compared to symmetric swapping of individuals used by Nettle).

The population we study consists of n individuals ($I_1 \dots I_n$) that can exchange m possible signals ($S_1 \dots S_m$). For each individual, the signal exchange is described by two matrices, P and Q , the first one defining production and the second comprehension. The elements p_{ij} of the production matrix P are the probabilities that a certain individual will refer to object i with signal j . Thus, the production matrix can be seen as a lookup table for generating signals associated with a specific object. The elements q_{ji} of the comprehension matrix Q are the probabilities that an individual will associate signal j with object i . As in Nowak's work, the ability of two individuals to communicate depends on their comprehension and production matrices. To measure this communication ability, a payoff function is introduced. The payoff $F(I_i, I_j)$ corresponds to the mutual understanding between the individuals I_i and I_j and is calculated given the four matrices Q and P of the two individuals (a more detailed description of the model can be found in Appendix A).

The evolution of language in this model is based on lexicon transmission from one generation to the next. In each iteration, the newly generated population is treated as students (listeners), while the old generation is treated as the teachers (speakers). Each student select its role models according to the learning mode of the simulation (the learning modes are described in the next section). Each teacher passes along its lookup table of objects (matrix P) and produces k signals associated with it. Note that in principle an object can be associated with several signals. The student receives the signals and keeps a record (represented by the association matrix A) of the number of times a specific object is associated with a specific signal. This

record is used at the end of the learning stage to build the comprehension and the production matrices of the new individual.

The maximum payoff for a P, Q pair is obtained for matrices P that have at least one 1 in every column, which means that each object is associated with only one signal. A more detailed analysis reveals that (1) this optimal state is an absorbing state of the system, and (2) an optimal learning mechanism will ultimately reach this state as there is an apparent bias in favor of one-to-one mappings between objects and words (Smith, 2002). When learning is complete and all the individuals have acquired their association matrices, the old generation is removed from the simulation (i.e. there is no overlap between generations), and the simulation calculates the Q and P matrices for the remaining individuals. The Q and P matrices of each individual are calculated using its association matrix A .

Three learning modes have been implemented and compared:

1. **Parental learning with fitness** (LINEAR SELECTION). In this mode, the offspring learn from their parents, with individuals that communicate well producing more offspring. Of the many possible fitness functions linking the number of offspring with the communication payoff we chose the linear fitness.
2. **Parental learning with no fitness** (DRIFT). In this mode, there is no advantage to individuals who communicate well.
3. **Role model** (SOCIAL SELECTION). In this mode, the members of a group acquire their language from a few selected “role model” individuals, singled out by their ability to communicate with their group members.

The migration component of our model is very similar to the original n -island approach of (Wright, 1931), and to the migration matrix approach of (Bodmer and Cavalli-Sforza, 1968). The island model assumes that individuals migrate from one subpopulation to another. In the initial state each sub-population consists of N individuals in M sub-populations. In each cycle, the migration rate M_{ij} determines the number of individuals relocating from population i to population j .

The migration rates, which are preset at the beginning of the simulation, may be the same for all sub-populations, or may vary according to some external factors (i.e. geographical distance). In the next section we address first the simple case in which the migration rates are fixed for all sub-populations; the more realistic case in which migration rates depend on geographical distances is treated later.

3 Results

We first analyzed the diversity of languages that evolved under each of the three fitness modes, by considering two kinds of indicators: (1) the *internal payoff*, which is the average payoff within a sub-population, and (2) the *external payoff*, which is the average payoff between individuals across the whole population. The migration matrix entries were constant and equal for all groups (in the more realistic case the migration matrix is merely symmetric). Figure 1(a) illustrates the time course of the internal and external payoffs across generations in the DRIFT mode. Notice that after the first 500 generations the internal payoff is higher than the external payoff. After 2000 generations, only one dominant language is left, at which point the internal payoff is equal to the external payoff. The same phenomenon was observed in the SOCIAL SELECTION, where, however, it took much longer for one dominant language to emerge. In comparison, in the LINEAR SELECTION mode (linear fitness), a diversity of languages is obtained under high internal and low external payoff.

To visualize the process of language development, one may consider the distance matrix whose elements d_{ij} correspond to the quality of understanding between any two individuals within the entire population:

$$d(ij) = 1 - F(I_i, I_j)/n \quad (1)$$

where $F(I_i, I_j)$ is the communication payoff between individuals I_i and I_j . Note that $d = 0$ means maximal understanding between two individuals, and $d = 1$ means no understanding. A simulation with 10 sub-populations of 50 individuals each yields a 500×500 distance matrix. The diagonal values are all equal to zero, because each individual understands itself perfectly.

Such a distance matrix captures the mutual understanding between individuals, regardless of their sub-population relation. To visualize the interaction between the individuals, one can represent each of them as a point in a multidimensional “understanding space”, in which the Euclidean distance between two points corresponds to the mutual understanding between the corresponding individuals. In such a representation, a sub-population that shares the same language will appear as a cluster, whose size is determined by its level of understanding (internal payoff value): a sub-population whose members have a high degree of mutual understanding will appear as a tight cluster.

For visualization purposes, we embed the understanding space into two dimensions, using a well-known technique (multidimensional scaling, or MDS (Shepard, 1980)). The resulting maps afford a certain insight into the pattern of sub-populations in each cycle. Figure 2 presents a sequence of snapshots of the understanding space, taken every 200 generations in the DRIFT, SOCIAL SELECTION, and LINEAR SELECTION

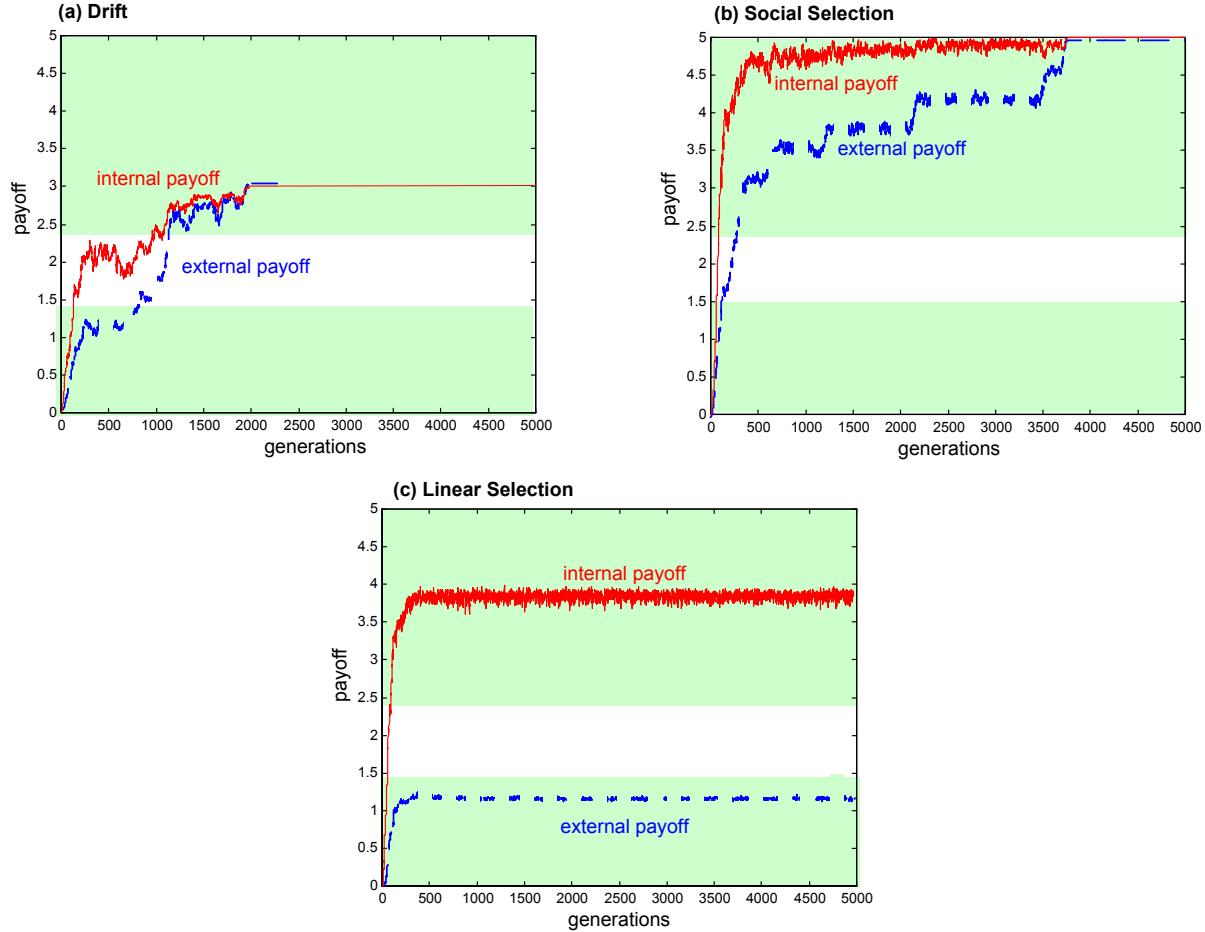


Figure 1: The internal payoff (solid line) and the external payoff (dashed line) in successive generations (the maximum possible payoff according to equation 2 is equal to the number of distinct signals $m = 5$). The upper band of the graph corresponds to a situation with one dominant language (all individuals understand each other regardless of their location). When the solid curve is in the upper band and the dashed one in the lower band, individuals within the sub-populations understand each other but not the members of other sub-populations (in other words, several languages co-exist). In this run we simulated 10 sub-populations of 50 individuals each. The migration rate between the islands was constant and equal to 0.01 (1% of migration per generation).

modes¹ In the DRIFT mode, the first several snapshots contain several clusters of languages; eventually, the clusters become unstable, that is, most of the languages become extinct and one language prevails. In comparison, in the SOCIAL SELECTION mode the system passes through a stage with several discrete clusters; after 2000 generations one dominant language remains. Likewise, in the LINEAR SELECTION mode, 10 sub-populations stabilize after 200 generations; after 800 generations only six of these are left.

In Figure 2, there is a large cluster in the DRIFT mode; in the SOCIAL SELECTION mode each of the much smaller clusters eventually converges to a single dot. These singleton clusters represent sub-populations with the most efficient language (recall that the size of a cluster is determined by its internal payoff value). In this sequence, all the clusters become singletons with the maximal internal payoff, which means that this fitness mode finds an optimally efficient language, but cannot converge on a stable multiple-language situation.

Figure 3 presents a sensitivity analysis of the three fitness modes for different migration rates. The only mechanism that remained stable against a higher rate of migration was the LINEAR SELECTION mode with linear fitness. Even in that mode, however, the system converged to a single language when the migration rate exceeded a critical value (about 4%).

Figure 4 illustrates a more realistic scenario, in which the sub-populations are randomly located on a 50×50 -cell and the migration rate between the groups is not fixed but determined geographically, as a Gaussian function of their distance. The MDS snapshots of the simulation were taken over three different iterations. As before, in this simulation too only the linear fitness mode can preserve the co-existence of several distinct languages. Furthermore, because of the geographical distance effect, sub-populations that are located close to each other are those whose languages are the first to merge.

4 Discussion

In the language evolution literature, one frequently encounters arguments based on the intuitive assumption that successful communication leads eventually to an increased chance of individual survival. This assumption occupies a prominent place in many models of language evolution, despite the scarcity of evidence supporting it. In this work, we addressed this issue by studying the evolutionary dynamics of language diversity with and without language-conferred fitness.

Our model, which combines the shared lexicon approach (Hurford, 1989; Nowak et al., 1999) with island migration methods (Nettle, 1999), allowed us to compare the relative contributions of language fitness and of population migration to language diversity. Although not intended to produce precise quantitative predictions, this approach makes it possible to study the general dynamics of the evolution of language

¹The circular cluster arrangement is an artifact that arises when MDS is applied to data that contain many equidistant pairs of points (Kruskal, 1977).

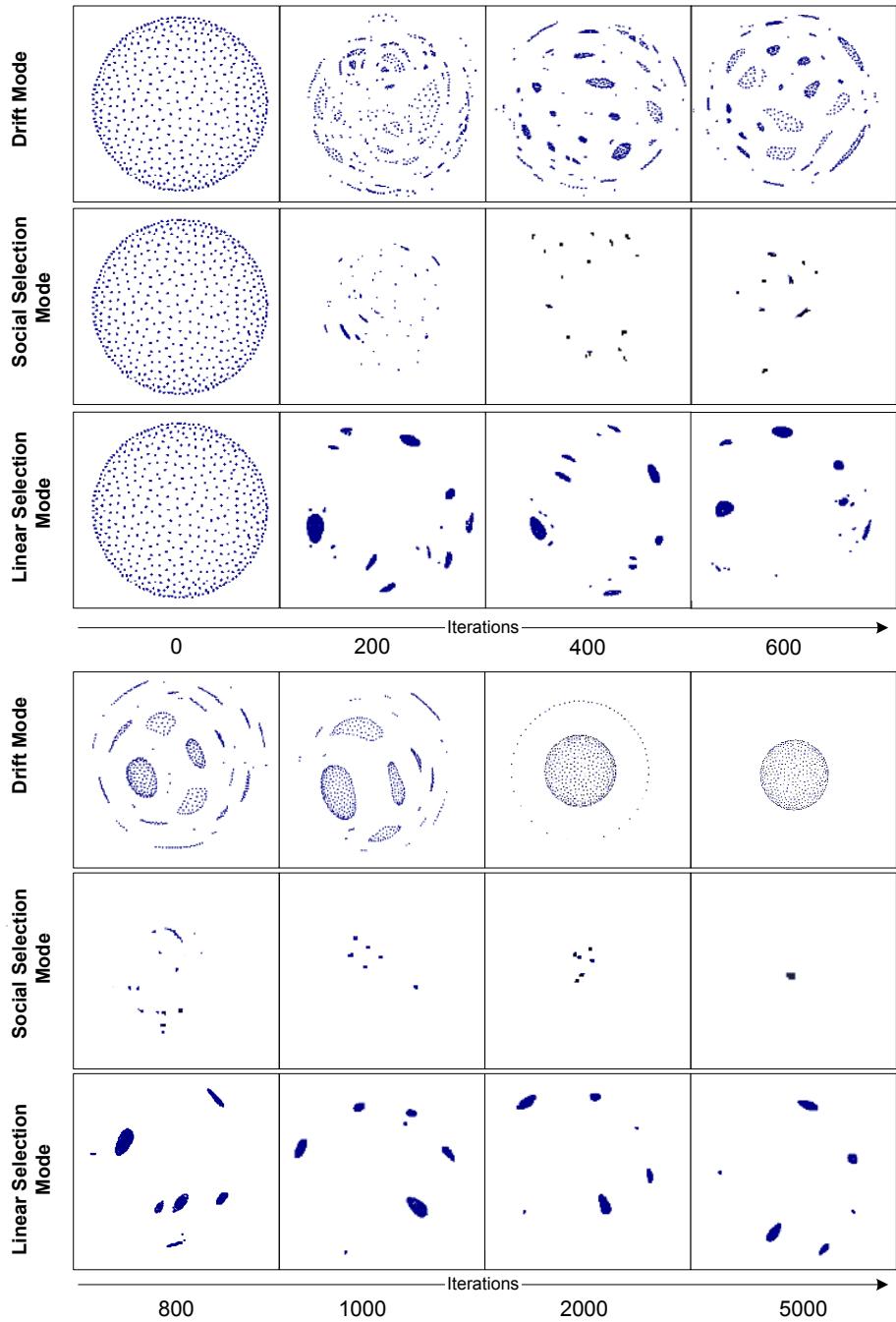


Figure 2: Multidimensional scaling (MDS) snapshots of the communication distance matrix under each of the three fitness modes. The circular arrangement of the clusters is a well-known artifact introduced by MDS (Kruskal, 1977), which is of no import in the present case (we are interested in the number of clusters and their shapes and not in their mutual arrangement).

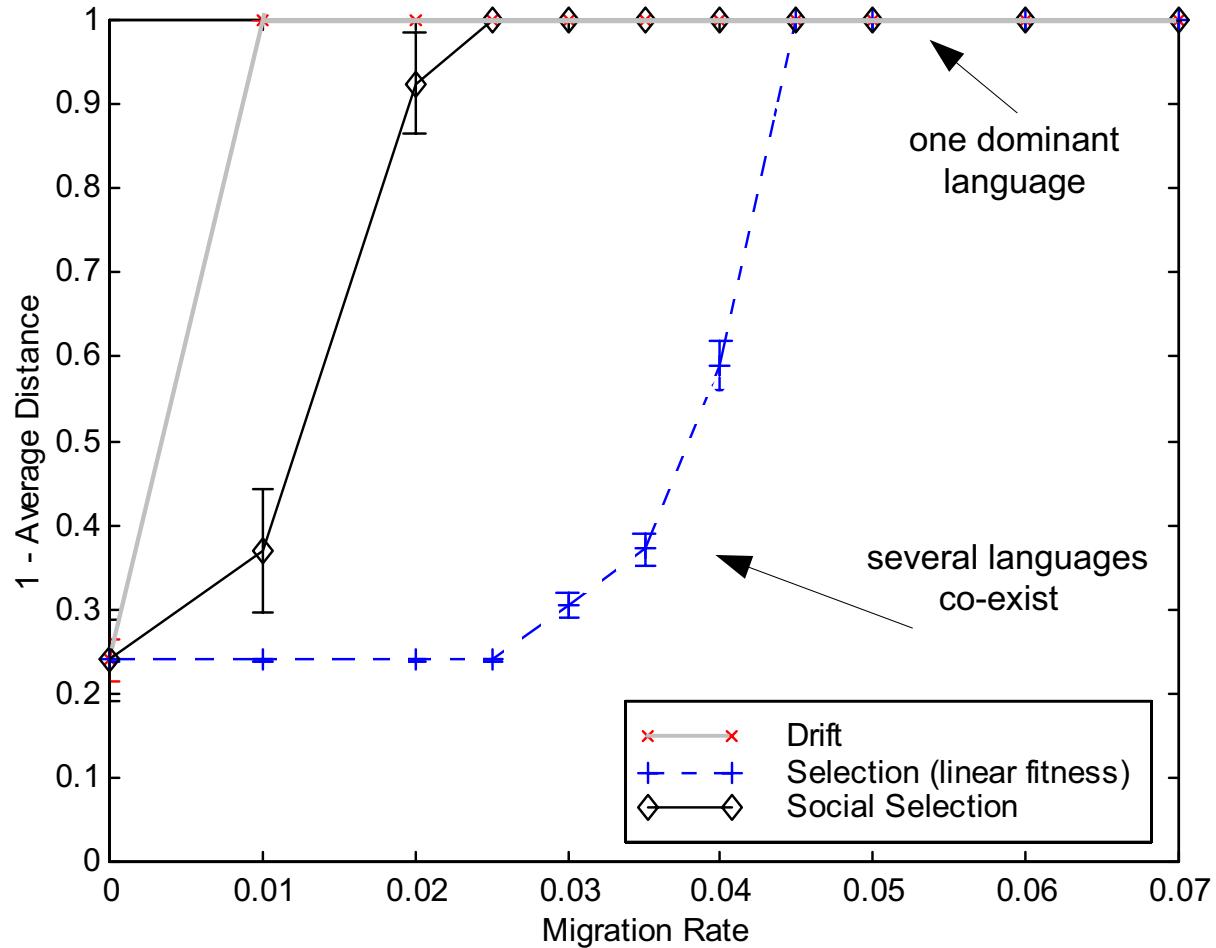


Figure 3: The dependence of linguistic diversity on migration rate under the three fitness modes (DRIFT, gray line; SOCIAL SELECTION, solid line; and LINEAR SELECTION, dashed line). The abscissa shows the migration rate and the ordinate – the internal-external payoff ratio, computed after 5000 generations by averaging the distance defined by equation 1. A ratio of 1 corresponds to a single dominant language; low ratios indicate high linguistic diversity.

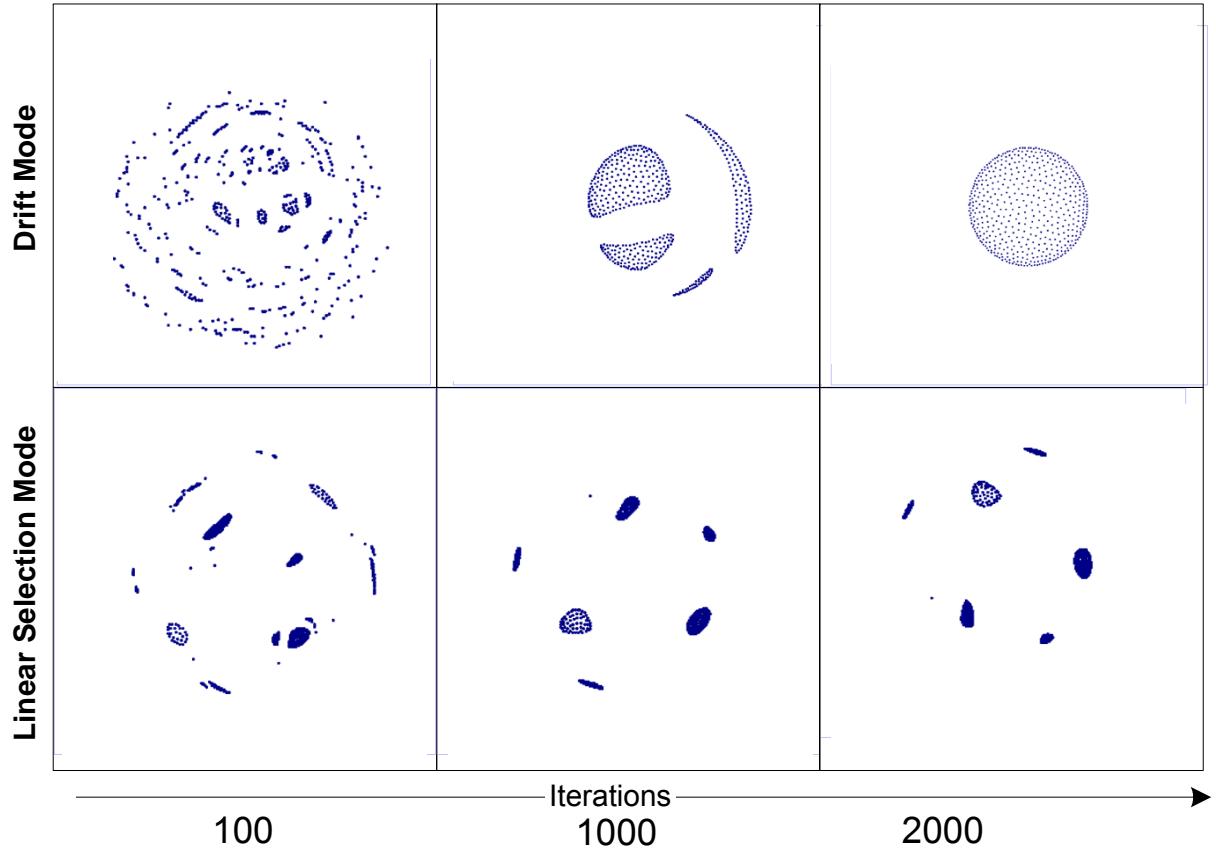


Figure 4: A spatial migration model. The migration rate between the groups is determined by a Gaussian function of their distance on a 50×50 grid. The figure shows the MDS snapshots of the state of the simulation, taken over three different iterations.

diversity. In particular, the behavior of the model in time and space offers some insights into the patterns of present-day linguistic diversity.

Our computational results differ from those of (Nettle, 1999) and Livingstone (Livingstone, 2000), most likely due to differences in the manner in which linear selection pressure and social effects have been implemented. In Nettle's model, these factors were not influenced by the ability of the individuals to communicate; in comparison, in the present study mutual understanding between individuals was the key factor shaping the evolutionary dynamics. Livingstone's study, in comparison, explored the effects of signal transmission but left out migration.

The results of our simulations suggest that the contribution of language to the fitness of individuals and the migration of individuals between sub-populations can indeed account for the evolution of language diversity. While fitness contributes positively to diversity, the contribution of migration is negative. We found that in the DRIFT and SOCIAL SELECTION conditions it takes much longer for the system to reach a steady state, and that even a very small migration rate suffices for one language to become dominant. In the LINEAR SELECTION case, when the fitness of individuals is assumed to be proportional to the success of their shared communication, the subpopulations stabilize quite rapidly to form several distinct languages. In this case, only relative high migration rates will force the system into a regime of one dominant language. Thus, given the continued presence of linguistics diversity and migration forces in the world, our results strongly suggest that fitness-based (linear) selection is at work in governing the evolutionary dynamics of language.

For each one of the cases (DRIFT, SOCIAL SELECTION and LINEAR SELECTION), a critical value of the migration rate can be identified, beyond which only one language can exist in the long run. This phenomenon offers a simple explanation for the transformation of social networks into linguistically bounded groups. One of the best known phenomena in linguistic diversity is the existence of small societies that in which distinct languages rather than continua of dialects are found. According to our findings, such sharp linguistic boundaries can arise when migration rates depend on geographical and sociological factors. Thus, in those locations where the migration rate is above the critical value, the languages will be fused into one; in locations that share a lower rate of migration, languages will remain divided. This situation was obtained in simulation 2 (see Figure 4), where in the linear selection mode, seven out of the initial 10 groups survived, as their migration rate was below the critical value. However, this phenomenon has not been observed in the absence of linear selection (i.e., in the DRIFT and SOCIAL SELECTION modes), as the migration rate among the groups exceeded the critical value, which in these modes is much lower. The effect of migration thus drives languages in these modes into extinction very quickly. We remark that the consistent and stable emergence either of one dominant language or of linguistic diversity for different values of a control parameter — the migration rate — has an intriguing analogy in solid-state physics: when certain materials are cooled rapidly, they crystallize heterogeneously (several crystals develop), while a slow gradual

annealing of the same material leads to a homogeneous crystallization.

Our analysis predicts that in places where the migration rate is beneath the critical value the diversity will be relatively high, while in places where it exceeds the critical value the level of diversity will be much lower. To examine this prediction, one may consider processes known to affect both migration rates and the level of language diversity. One such factor is *ecological risk* suggested by Nettle (Nettle, 1999), which is defined as the probability of a household to face a temporary shortfall in food production.² Nettle showed that in places where the ecological risk factor (and therefore the migration rate) is low the level of diversity is high, and vice versa. Another relevant factor is social networks: where such networks exist the level of diversity is low (e.g., Niger), while in places where the social networks are less developed (e.g., Papua New Guinea) the level of linguistic diversity is much higher. Bringing additional factors, such as means of transportation, mobility and average distance among sub-populations, into the picture should lead to a better understanding of the relation between migration rates and language diversity.

Another prediction of our model concerns the development of language diversity over time. Specifically, our model predicts a non-constant rate of language evolution, where the appearance of new languages is very rapid (the first stage of language evolution takes very few iterations), and the extinction of languages is slow (see Figure 2). Dixon (Dixon, 1997) has proposed a controversial model (cf. (Joseph, 2001)) of the development of language over time, building on ideas from biological evolution (Eldredge and Gould, 1972). In his “punctuated equilibrium” model, language evolution undergoes periods of rapid expansion, during which many languages evolve, interspersed with long periods of near-equilibrium, during which languages diffuse and converge, as societies interact, intermarry, fission and fuse. In comparison to Dixon’s, our approach needs not postulate *ad hoc* discrete evolutionary stages: the variable rate of language evolution, spread and disappearance emerges from the dynamics of the underlying processes of differentiation and drift.

References

- Batali, J. (1998). *Computational simulations of the emergence of grammar*, pages 405–426. Cambridge University Press, Cambridge.
- Bodmer, W. F. and Cavalli-Sforza, L. L. (1968). A migration matrix model for the study of random genetic drift. *Genetics*, 59:565–592.
- Briscoe, T. (2003). Coevolution of the language faculty and language(s) with decorrelated encodings. In *OUP language evolution volume*, page see this volume.

²People tend to migrate more in regions with higher ecological risk.

- Cangelosi, A. (2001). Evolution of communication and language using signals, symbols, and words. *IEEE Transactions on Evolutionary Computation*, 5(2):93–101.
- Dixon, R. (1997). *The Rise and Fall of Languages*. Cambridge University Press, Cambridge, MA.
- Eldredge, N. and Gould, S. J. (1972). Punctuated equilibria: an alternative to phyletic gradualism models of paleobiology. In *Models in Paleobiology*, pages 82–115. Freeman, Cooper, San Francisco, CA.
- Grimes, P. (2001). *data from Ethnologue: Languages of the World, 14th Edition*. SIL International.
- Hurford, J. (1989). Biological evolution of the saussurean sign as a component of the language acquisition device. *Lingua*, 77(2):187–222.
- Joseph, B. (2001). Review of R.M.W. Dixon, the rise and fall of languages. *Journal of Linguistics*, 37(1).
- Kirby, S. (2000). *Syntax without Natural Selection: How compositionality emerges from vocabulary in a population of learners.*, pages 303–323. Cambridge University Press.
- Kirby, S. (2002). *Learning, Bottlenecks and the Evolution of Recursive Syntax*, chapter 6. Cambridge University Press.
- Komarova, N. L. and Nowak, M. A. (2001). The evolutionary dynamics of the lexical matrix. *Bull. Math. Biol.*, 63(3):451–485.
- Kruskal, J. B. (1977). The relationship between multidimensional scaling and clustering. In Ryzin, J. V., editor, *Classification and clustering*, pages 17–44. Academic Press, New York.
- Livingstone, D. (2000). Neutral evolution and linguistic diversity. Technical Report 9, University of Paisley, Computing and Information Systems.
- Livingstone, D. and Fyfe, C. (2000). Modeling language-physiology coevolution. In Chris Knight, J. R. H. and Studdert-Kennedy, M., editors, *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*. Cambridge University Press, Cambridge.
- Nettle, D. (1999). *Linguistic Diversity*. Oxford University press.
- Nichols, J. (1992). *Linguistics Diversity in Space and Time*. University of Chicago Press.
- Nowak, M. A. and Krakauer, D. (1999). The evolution of language. *Proceedings of the National Academy of Sciences of the United States of America*, 96(14):8028–8033.
- Nowak, M. A., Plotkin, J. B., and Jansen, V. A. A. (2000). The evolution of syntactic communication. *Nature*, 404:495–498.

- Nowak, M. A., Plotkin, J. B., and Krakauer, D. C. (1999). The evolutionary language game. *Journal of Theoretical Biology*, 200:147–162.
- Oliphant, M. (1999). The learning barrier: Moving from innate to learned systems of communication. *Adaptive Behavior*, 7(3/4).
- Oliphant, M. and Batali, J. (1997). Learning and the emergence of coordinated communication. *The newsletter of the Center for Research in Language*, 11(1).
- Reggia, J. A., Schulz, R., Wilkinson, G., and Uriagereka, J. (2001). Conditions enabling the evolution of inter-agent signaling in an artificial world. *Artificial Life*, 7(1):3–32.
- Shepard, R. N. (1980). Multidimensional scaling, tree-fitting, and clustering. *Science*, 210:390–397.
- Smith, K. (2001). The importance of rapid cultural convergence in the evolution of learned symbolic communication. In *ECAL01*, Lectures Notes in Computer Science, pages 637–640, Prague. Springer.
- Smith, K. (2002). The cultural evolution of communication in a population of neural networks. *Connection Science*, 14(1):65–84. eprint-ID: taal00000004.
- Wright, S. (1931). Evolution in mendelian populations. *Genetics*, 16:97–159.

5 APPENDIX A

The population we study consists of n individuals ($I_1 \dots I_n$) that can exchange m possible signals ($S_1 \dots S_m$). For each individual, the signal exchange is described by two matrices, P and Q , the first one defining production and the second comprehension. The elements p_{ij} of the production matrix P are the probabilities that a certain individual will refer to object i with signal j ; each row in the matrix P sums to 1. The elements q_{ji} of the comprehension matrix Q are the probabilities that an individual will associate signal j with object i ; here too each row sums to 1. The payoff value corresponds to the mutual understanding between the two individuals; an individual's average payoff is calculated by summing its payoff with respect to each of its group members, divided by the group size (see equation 2).

$$\begin{aligned} F(I_k, I_l) &= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^m \left(p_{ij}^{(k)} q_{ji}^{(l)} + p_{ij}^{(l)} q_{ji}^{(k)} \right) \\ F(I_u) &= \frac{1}{n-1} \sum_{v \neq u} F(I_u, I_v) \end{aligned} \quad (2)$$

The evolution of language in this model is based on lexicon transmission from one generation to the next. Individual I_k learns from I_l by sampling the responses of I_l to every object, This is incorporated in the

model by sequentially scanning the rows of I_l 's production matrix (each row corresponds to a different object). For each row, the agent produces k random signals according to the probabilities listed. The listener I_k undergoes a “learning phase” in which it updates its association matrix A ; this is an $n \times m$ matrix whose entries a_{ij} specify how often the individual has observed the object i being referred to by the signal j . The Q and P matrices are both derived from A according to equation 3.

The three learning modes we considered are: Parental learning with fitness, (LINEAR SELECTION). Parental learning with no fitness (DRIFT), and Role model (SOCIAL SELECTION).

$$p_{ij} = \frac{a_{ij}}{\sum_{l=1..m} a_{il}} \quad q_{ji} = \frac{a_{ij}}{\sum_{l=1..n} a_{lj}} \quad (3)$$

The migration component of our model is very similar to the original n -island approach of (Wright, 1931), and to the migration matrix approach of (Bodmer and Cavalli-Sforza, 1968). The former assumes that individuals migrate from one sub-population to another with the same rate, which means that the distances between subpopulations are not taken into account (cf. matrix M_1 in equation 4 and Figure 5). In the latter approach, the migration rates between sub-populations can be different, e.g., depending on distance (cf. matrix M_2 in equation 4 and Figure 5). In our implementation, each sub-population consisted of N individuals. In each cycle, the migration rate determined the number of individuals relocating from one island to another; a migration matrix controlled the migration rate between the sub-populations.

$$M_1 = \begin{pmatrix} 0 & m & m & m \\ m & 0 & m & m \\ m & m & 0 & m \\ m & m & m & 0 \end{pmatrix} \quad M_2 = \begin{pmatrix} 0 & m & \frac{m}{2} & \frac{m}{4} \\ m & 0 & m & \frac{m}{2} \\ \frac{m}{2} & m & 0 & m \\ \frac{m}{4} & \frac{m}{2} & m & 0 \end{pmatrix} \quad (4)$$

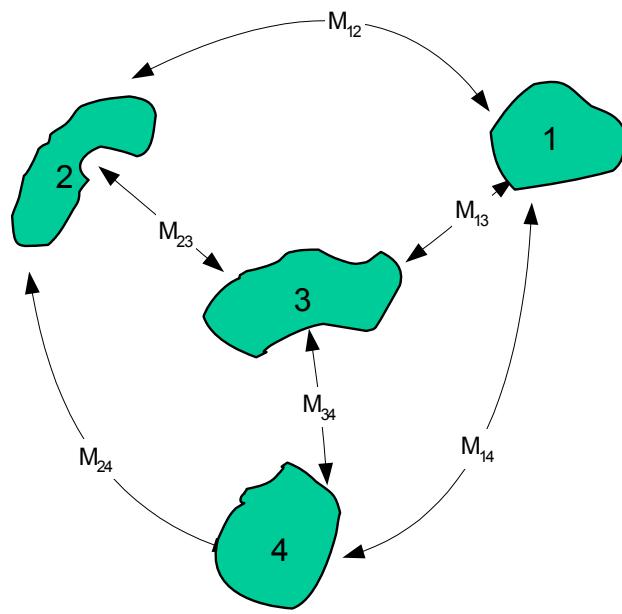


Figure 5: The island migration model (general case). The island model assumes that individuals migrate from one sub-population to another with different migration rate values. Initially, there are N individuals in each of the M sub-populations. In every cycle, the migration rate M_{ij} determines the number of individuals relocating from population i to population j . The migration rates M_{ij} can be arranged in a migration matrix.