J. theor. Biol. (1999) **200,** 147–162
Article No. jtbi.1999.0981, available online at http://www.idealibrary.com on IDEAL®

The Evolutionary Language Game

MARTIN A. NOWAK*, JOSHUA B. PLOTKIN AND DAVID C. KRAKAUER

Institute for Advanced Study, Olden Lane, Princeton, NJ 08540, *U.S.A.*

(*Received on* 25 *February* 1999, *Accepted in revised form on* 2 *June* 1999)

We explore how evolutionary game dynamics have to be modified to accomodate a mathematical framework for the evolution of language. In particular, we are interested in the evolution of vocabulary, that is associations between signals and objects. We assume that successful communication contributes to biological fitness: individuals who communicate well leave more offspring. Children inherit from their parents a strategy for language learning (*a language acquisition device*). We consider three mechanisms whereby language is passed from one generation to the next: (i) parental learning: children learn the language of their parents; (ii) role model learning: children learn the language of individuals with a high payoff; and (iii) random learning: children learn the language of randomly chosen individuals. We show that parental and role model learning outperform random learning. Then we introduce mistakes in language learning and study how this process changes language over time. Mistakes increase the overall efficacy of parental and role model learning: in a world with errors evolutionary adaptation is more efficient. Our model also provides a simple explanation why homonomy is common while synonymy is rare.

1. Introduction

A unity of design and a diversity of implementations are defining features of evolved traits. Common descent provides a set of shared characteristics upon which local adaptation and ontogenetic mechanisms impress their modifications. For a small set of species learning has opened up the possibility of almost limitless variety, to the extent that the underlying uniformity has been obscured. This is particularly true for human language and to a lesser extent, for signal evolution in animals.

Humans from different language communities employ different sounds to denote the same object, action or property. Among the song birds, dialects are highly variable and change within an individual over the course of its life (Lemon, 1975; Catchpole *et al*., 1984; Balaban, 1988; Kroodsma & Konishi, 1991). The common structures of human and some animal signals have in part been elucidated (Chomsky, 1965, 1972, 1975, 1980; Frisch 1967; Marler, 1970; Marler & Pickert, 1984; Marler *et al*., 1988). It is now thought that features of human language are "universal" and constant across language communities: the principles of language are shared and have evolved, while the parameters are free to change through learning (Pinker & Bloom, 1990; Pinker, 1994). This begs the questions, why learn signals at all and why not simply inherit them? Why when these signals have been learnt do they so often change? Which members of a community should serve as models for imitation? And which components of a strategy are inherited and which ones learnt?

---

* Author to whom correspondence should be addressed. E-mail: kathei@ias.edu

0022–5193/99/018147 + 16 $30.00/0

Imitation like language is rare in nature (Galef, 1988). Since language learning in humans requires imitation, a part of the explanation for language's rarity might be ascribed to the difficulties of imitation. The problem is in part evolving a means of associating an action with a goal and a reward without experiencing it oneself. Furthermore, which individual of a population is best to imitate? A safe place to begin is with ones parents. Human infants of a few weeks in age, and even newborns, show a tendency to imitate the facial gestures of those around them (Meltzoff, 1983; Meltzoff & Moore, 1994). Imitation is central to language acquisition but there is more to learning a language than simple imitation. Language acquisition is not the same as learning a skill through guided trial and error because it requires the formation of a shared representation among signallers. Linguistic signals are referential and hence imitation need not imply understanding. Imitation of a signal when coupled to an inappropriate behaviour will be selected against.

There are numerous publications on the evolution of human language ranging from studies of primate or other animal behaviour (Smith, 1977; Seyfarth *et al.*, 1980; Burling, 1989; Cheney & Seyfarth, 1990; Poran & Coss, 1990; Greenfield, 1991; Hauser, 1996; Dunbar, 1997; Bradbury & Vehrencamp, 1998), to the development of language in young children (Brown, 1973; Anisfield, 1984; Sachs, 1985; Newport, 1990; Hurford, 1991; Bates, 1992), the genetic and anatomic correlates of language competence (Wittelson & Pallie, 1973; Lieberman, 1984, 1991; Newport, 1990; Nobre *et al.*, 1994; Deacon, 1997) and the diversity of existing human languages (Greenberg, 1971; Cavalli-Sforza & Cavalli-Sforza, 1995; Cavalli-Sforza, 1997). Maynard Smith & Szathmary (1995) described human language as a major transition in evolution (see also Szathmary & Maynard Smith, 1995). Bickerton (1990) argues that this transition is likely to have been abrupt, while Pinker (1994) has argued that a gradual process is likely to have been sufficient. Most studies of language evolution do not use mathematical models. There are some exceptions (Aoki & Feldman, 1987, 1989; Hurford, 1991; MacLennan, 1992; Hashimoto & Ikegami, 1995; Hutchins & Hazelhurst, 1995; Oliphant, 1996; Parisi, 1997; Palacious, 1998; Steels, 1998; Briscoe, 1999; see also Hurford *et al.*, 1999). In an early theoretical paper, Hurford (1989) discussed a mathematical model for the evolution of Saussarian signs (Saussure, 1916). Hurford's model is similar to our approach here and also uses two matrices to describe communication. Cavalli-Sforza & Feldman (1981) and Boyd & Richardson (1985) provide alternative mathematical frameworks for cultural transmission in evolution.

Our aim here is to expand the standard assumptions of evolutionary game theory and develop a general, but simple mathematical model for understanding the evolution of language. In particular, we are interested in the evolution of a common vocabulary (that is a specific association between signals and objects). We do not explore the evolution of syntax or grammar in this paper. We extend an approach taken by Nowak & Krakauer (1999) and study in more detail how language is transmitted from one generation to the next. In Nowak & Krakauer (1999), we simply assumed that offspring adopt the language of their parents; here we specify learning mechanisms. Section 2 defines the model. Section 3 formalizes language learning. Section 4 outlines the population dynamics and discusses three mechanisms whereby language is transmitted from one generation to the next. Section 5 takes advantage of mistakes. Section 6 shows results on competition between the mechanisms of language transmission. Section 7 concludes. Appendix A classifies some binary matrices, which are the absorbing states of our stochastic evolutionary processes. Appendix B describes a paradoxical result: two individuals may understand each other better than themselves.

2. Defining the Model

Consider a group of individuals (animals or early hominids) able to produce a number of signals (sounds). Information shall be transferred about a number of objects; we use "objects" (or concepts) in an extended sense to include other individuals, other animals, plants, inanimate objects, actions or events (all things that can be referred to). Suppose there are m signals and

n objects. The matrix P contains the entries p_{ij}, which denote the probability that for a speaker object i is associated with sound j. P is an $n \times m$ matrix whose rows sum to 1:

$$\sum_{j=1}^{m} p_{ij} = 1. \tag{1}$$

The matrix Q contains the entries q_{ji}, which denote the probability that for a listener sound j is associated with the object i. Q is an $m \times n$ matrix whose rows sum to 1:

$$\sum_{i=1}^{n} q_{ji} = 1. \tag{2}$$

P describes speaking, whereas Q describes listening (comprehension). We call P the "active matrix" and Q the "passive matrix".

Imagine two individuals, I_1 and I_2, who use languages L_1 (given by P_1 and Q_1) and L_2 (given by P_2 and Q_2). For individual I_1, $p_{ij}^{(1)}$ denotes the probability of making sound j when seeing object i, whereas $q_{ji}^{(1)}$ denotes the probability of inferring object i when hearing sound j. For individual I_2, these probabilities are given by $p_{ij}^{(2)}$ and $q_{ji}^{(2)}$. Suppose I_1 sees object i and signals, then I_2 will infer object i with probability $\sum_{j=1}^{m} p_{ij}^{(1)} q_{ji}^{(2)}$. A measure of I_1's ability to convey information to I_2 is given by summing this probability over all objects, n: $\sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij}^{(1)} q_{ji}^{(2)}$. The overall payoff for communication between I_1 and I_2 is taken as the sum of I_1's ability to convey information to I_2, and I_2's ability to convey information to I_1. Thus,

$$F(L_1, L_2) = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{m} (p_{ij}^{(1)} q_{ji}^{(2)} + p_{ij}^{(2)} q_{ji}^{(1)}). \tag{3}$$

In this equation, both individuals are treated once as a listener and a speaker, which leads to the intrinsic symmetry of the language game: $F(L_1, L_2) = F(L_2, \mathrm{L}_1)$. Language L_1 obtains from L_2 the same payoff as L_2 from L_1. If two individuals use the same language, L, the payoff is

$$F(L, L) = \sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} q_{ji}. \tag{4}$$

The passive matrix, Q, can be treated as completely independent of P. In this case, it is not taken for granted that an individual who signals "leopard" when seeing a leopard will also conceive of a leopard when hearing "leopard"; this symbolic association is something which must be formed by the evolutionary dynamics of the system. Alternatively, we could assume that individuals have an internal mechanism that links their active and passive matrices.

2.1. SOME COMMENTS ON P AND Q

2.1.1. *The best possible Q*

An interesting question is: for a given P what is the optimum Q that maximizes the payoff $F(L, L)$? We find that the optimum Q is always a binary matrix such that $q_{ji} = 1$ if p_{ij} is the largest entry in a column of P; all other entries of Q are 0. (In case there are 2 or more equally large entries in a column of P one may be chosen at random.) Thus for a given matrix P, the maximum payoff is given by

$$F_{max}(L, L) = \sum_{j=1}^{m} p_{*j}, \tag{5}$$

where p_{*j} is the maximum entry in column j of the P matrix.

2.1.2. *The best possible P and Q*

The maximum payoff for a P and Q pair is obtained for binary active matrices P that have at least one 1 in every column (if $n \geqslant m$) or in every row (if $n \leqslant m$), while Q is constructed as described in Section 2.1.1. For $n = m$ the maximum payoff is obtained if P has exactly one 1 in every row and column (while every other entry is 0) and $Q = P^{\mathrm{T}}$ (Q is the transposed matrix of P). Such a matrix is called a "permutation matrix". In general, the maximum payoff is given by

$$F_{max}(L, L) = \min\{m, n\}. \tag{6}$$

3. Learning a Language

We will assume that individuals acquire a language (a lexicon) by observing and imitating other individuals. Specifically, we assume that each individual undergoes a "learning phase" during which it constructs an "association matrix", A. A is an $n \times m$ matrix whose entries, a_{ij}, specify how often an individual has observed one or several other individuals referring to object i by producing signal j. Therefore, the entries of A are nonnegative integers. The active and passive matrices are then derived from the association matrix by normalizing rows and columns, respectively:

$$p_{ij} = a_{ij} \Bigg/ \left(\sum_{l=1}^{m} a_{il} \right), \qquad q_{ji} = a_{ij} \Bigg/ \left(\sum_{l=1}^{n} a_{lj} \right). \tag{7}$$

Suppose individual I_2 learns from I_1. We assume that I_2 samples the responses of I_1 to every object k times. In the limit $k \to \infty$, I_2 will exactly reproduce I_1's active matrix. For $k = 1$, I_2's association matrix, A_2, will be binary and $P_2 = A_2$. Thus, $k = 1$ leads immediately to binary active matrices.

Assume that all individuals of a population speak the same language. Consider the stochastic process

$$P_0 \to A_1 \to P_1 \to A_2 \to P_2 \to \cdots. \tag{8}$$

A given active matrix, P_i, is sampled to give an association matrix, A_{i+1}, which is then converted into P_{i+1} (and Q_{i+1}) according to eqn (7). The language changes over time because language learning (constructing A from P) is a probabilistic process. Without errors in language learning the process will ultimately reach an absorbing state. The absorbing states are given by the set of all binary P matrices where each object is associated with only one signal (but a given signal may refer to several objects).

3.1. HOMONOMY AND SYNONYMY

There is a general understanding among language scholars that homonomy is plentiful in language (any common word in the dictionary will have several meanings) whereas synonymy is rare (it is hard to find two words that have exactly the same meaning). Interestingly, this observation is a direct consequence of the language learning mechanism described by eqn (9).

Homonomy refers to the case where we have more than one non-zero entries in a column of the active matrix of the language: thus two different objects (concepts) can be associated with the same word. This situation is stable. Consider the active matrix

$$P = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}.$$

If sampled by the child, it will lead to the association matrix

$$A = \begin{pmatrix} k & 0 \\ k & 0 \end{pmatrix},$$

which again leads to the same active matrix. Homonomy is an absorbing state of our stochastic process.

Synonymy refers to the case where we have more than one non-zero entries in a row of the active matrix: thus, the same object (concept) is associated with two different signals. This situation is not stable. Consider the active matrix

$$P = (0.5 \quad 0.5).$$

Binomial sampling will lead to an association matrix which is slightly asymmetric resulting in an asymmetric new P matrix, which most likely gives rise to an even more asymmetric A matrix. The only stable solutions are given by

$$P = (1 \quad 0) \quad \text{or} \quad P = (0 \quad 1),$$

where synonymy has disappeared.

4. Population Dynamics

Let us now consider a population of N individuals speaking languages L_1 to L_N. Each individual talks to every other individual and receives a total payoff,

$$F_I = \sum_J F(L_I, L_J), \tag{9}$$

where $J = 1, \ldots, N$, but $J \neq I$. We explore three different mechanisms for transmitting the language from one generation to the next. These

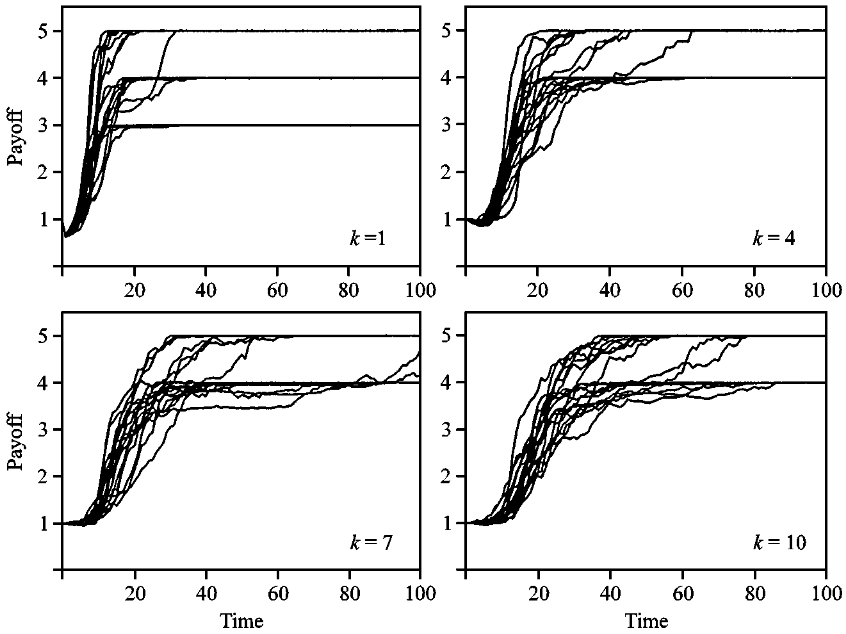

FIG. 1. A common language evolves if individuals with a high payoff in the language game produce more offspring and if children learn the language of their parents. There are $N = 100$ individuals communicating about $n = 5$ objects using (up to) $m = 5$ signals. There are discrete generations. In one generation each individual is talking to every other individual. The payoffs of all interactions are summed up. Individuals produce offspring for the next generation proportional to their total payoff. Offspring sample the active matrix from their parents thereby forming their own association matrix, A, between objects and signs. For each object, k responses are sampled. Individuals derive their P and Q from their A. Initially, every individual has a random A; a_{ij} are taken from a uniform distribution on the interval (0, 1). The four panels show 20 runs each for $k = 1, 4, 7, 10$. All simulations eventually settle to binary P matrices achieving payoffs 6, 8, or 10. The case $k = 1$ has the fastest convergence to a common language but reaches on average a slightly lower final payoff than $k = 4$, 7, or 10.

mechanisms represent different strategies for language acquisition.

4.1. PARENTAL LEARNING

First, we assume that children learn the language of their parents. Figure 1 shows results of computer simulations for this case. There is a population of $N = 100$ individuals using (up to) $m = 5$ signals to communicate $n = 5$ objects. Initially, all individuals have random matrices. In each generation, every individual "talks" to all others and the payoffs are summed up. The total payoff for each individual is given by eqn (10). For the next generations, offspring are produced proportional to the total payoff of an individual: the chance that a particular individual arises from individual I is given by $F_I/\sum_J F_J$. There is no sexual reproduction. Each individual has one parent. The individuals of the new generation learn the language of their parents by sampling their responses to specific objects. Each individual forms its association matrix, A, by recording k responses of its parent to each object.

Figure 1 shows 20 runs for four different values of k: 1, 4, 7, 10. Starting from a random initial condition, the population converges to a common language. The evolutionary optimum is reached if every word is associated with exactly one signal (and vice versa). This optimum (which results in a payoff of 5 points) is reached often, but not always. Sometimes the population settles for a language where two different signals are associated with one object, whereas another signal is used for two different objects. This suboptimum case achieves payoff 4. Payoff 3 is also reached occasionally. In this case, either the same signal is used for three different objects, or there are two pairs of objects each associated with one signal. There is no significant difference in the average final outcome for different values of k. For higher values of k, however, the population takes a longer time to converge to a common language.

4.2. ROLE MODEL LEARNING

A different concept is that language may confer a certain reputation in the group, which can be associated with an increased chance of imitation by members of the next generation. In this case, it seems at first sight unimportant whether successful individuals also produce more genetic offspring; they transmit their language culturally. Therefore, we consider a scenario in which individuals with a high payoff have a higher chance of being imitated. The probability that a member of the new generation samples the active matrix of individual I of the previous generation is given by $F_I/\sum_J F_J$. Each individual samples K individuals of the previous generation. For $K = 1$, this process leads to the identical algorithm as for the case described in Section 4.1. Thus, parental learning and role model learning can be described using the same formalism. The equivalence breaks down for $K > 1$; there can be only one genetic parent (without sexual reproduction), but several cultural parents. (Note the difference between K and k: the parameter K denotes the number of individuals that are sampled, while k counts how often a given individual is sampled.)

Figure 2 shows computer simulations for $K = 1, 4, 7, 10$. For $K = 1$, we have rapid convergence; the average payoff is equivalent to that found in Fig. 1. For $K > 1$, convergence takes much longer, but the average payoff is significantly higher than for $K = 1$. There is, however, no significant difference between $K = 4$, 7 and 10.

4.3. RANDOM LEARNING

Finally, we consider the situation where children learn the language of random members of the previous generation, irrespective of their payoff. There is no (direct) reward for high payoff and consequently no selection on improving language efficiency. Nevertheless, the resulting

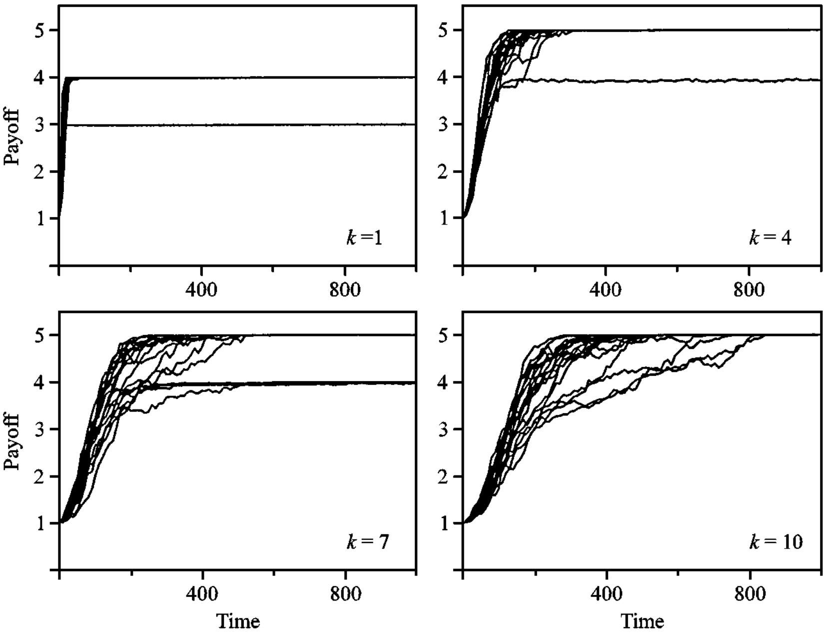

FIG. 2. A common language evolves if children learn their language from individuals in the previous generation who have a high payoff. There are $N = 100$ individuals communicating about $n = 5$ objects using (up to) $m = 5$ signals. There are discrete generations. In one generation each individual is talking to every other individual. The payoffs of all interactions are summed up. Each individual of the next generation choses K individuals of the previous generation and samples their responses (1 response per object per individual is sampled). P and Q are derived from A. The probability that an individual is chosen for being sampled is proportional to its payoff in the language game. Thus successful individuals have a higher chance of passing on their language. The four panels show 20 runs each for $K = 1, 4, 7, 10$. All simulations eventually settle to binary P matrices achieving payoffs 6, 8, or 10. The case $K = 1$ has the fastest convergence to a common language but reaches on average a slightly lower final payoff. $K = 10$ has slowest convergence, but reaches this maximum payoff in 20 of 20 cases. This is, however, a coincidence. In simulations using 100 runs per K-value, there is no significant difference between $K = 4$, 7 or 10, although $K = 1$ performs clearly worse.

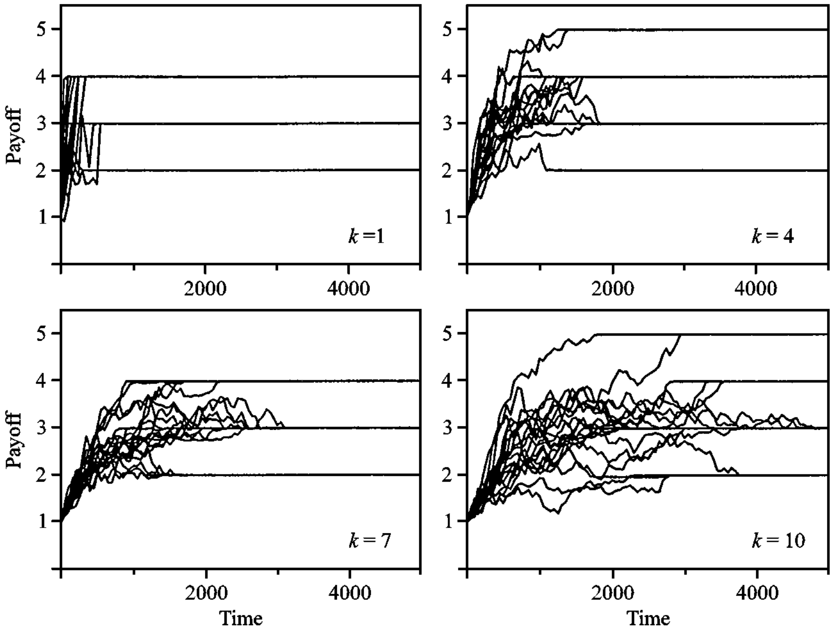

FIG. 3. A common language also evolves if children learn their language from randomly chosen individuals of the previous generation irrespective of any payoff. In this case, however, there is no reward for efficient communication. The active matrix, P, is not at all optimized but drifts to an arbitrary binary matrix. (Note that the binary matrices are the absorbing states of this random process). All parameter values are as for Fig. 2, but each individual samples K *randomly* chosen individuals of the previous generation. Note that convergence is much slower than in Fig. 2 and leads on average to lower final payoff.

stochastic process can lead to fairly efficient communication systems, because the absorbing states of the high-dimensional "random walk" are random binary matrices, which on average achieve a higher payoff than random matrices where the entries are numbers between 0 and 1. In Appendix A, we discuss the combinatorics of these binary matrices and calculate their average payoff as well as the fraction of random binary matrices that obtain maximum payoff: as m gets large compared to n, the average fitness approaches the maximum fitness and the fraction of binary matrices that achieve maximum fitness converges to 1.

Figure 3 shows computer simulations for the case where children learn their language from K random individuals of the previous generation, irrespective of their payoff. Convergence to a common language is slower for larger values of K and in general, the convergence is slower than in Figs 1 and 2. For all values of K, however, the same average payoff is achieved. The distribution of payoff values simply reflects the statistics of random binary matrices. Note that the mathematical algorithm for modelling random transmission in Fig. 3 does not require any payoff evaluation as the payoff achieved in the language game is totally irrelevant here.

5. Take Advantage of Mistakes

Let us add another fundamental feature to our model: language acquisition should be error-prone. Assume that with probability, ρ, a learning individual mistakes the response of another individual for a randomly chosen response. Thus with probability $1 - \rho$ the learning individual makes the correct entry into its association matrix, A, while with probability ρ it makes a random entry for this particular object. (Here we consider errors in language learning; for the effect of errors in communication we refer to Nowak & Krakauer, 1999.)

Figure 4 shows the same simulation as Fig. 1 (for $k = 1$) but with four different noise levels ($\rho = 0.0001, 0.001, 0.01$ and 0.1). For low noise levels ($\rho = 0.0001$ and 0.001) the system behaves similar to Fig. 1. For $\rho = 0.01$, however, all

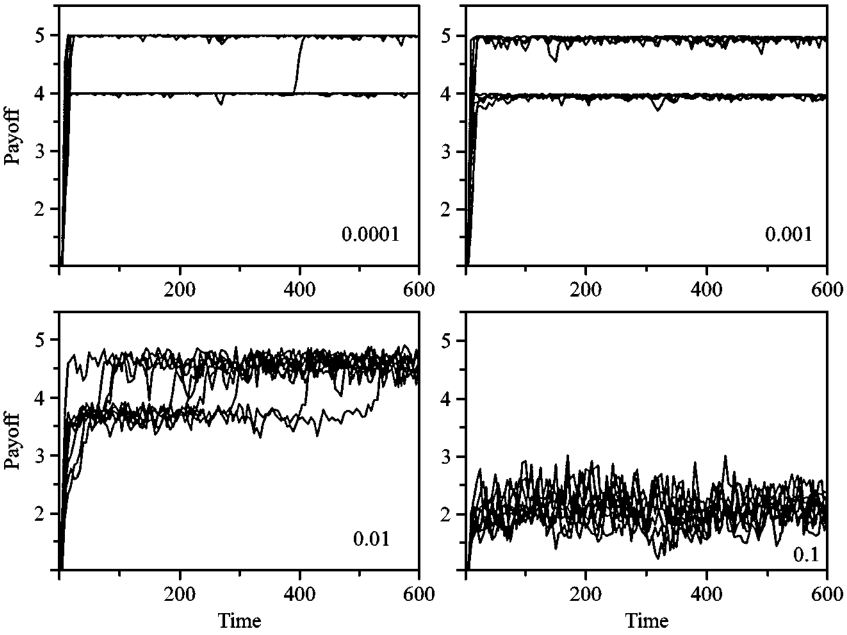

FIG. 4. Errors during language acquisation favour the evolution of more efficient languages. Here the simulations of Fig. 1 with $k = 1$ are repeated but individuals make mistakes during language learning: with probability ρ the response of another individual to a given object is mistaken for a randomly chosen response. For small levels of noise, $\rho = 0.0001$ and 0.001, there is not much difference compared to the case without errors. For $\rho = 0.01$ all simulations converge to a global optimum. For $\rho = 0.1$ the system does not evolve efficient communication. Thus there seems to be an optimum error rate.

simulations converge to the maximum payoff where each signal is associated with exactly one object. Errors prevent the system from getting trapped in sub-optimum situations where the same signal would be used for two or more objects. For high probabilities of mistakes ($\rho = 0.1$) the optimization ability of the system is impaired.

Thus, without noise the system converges to one of the absorbing states (given by all random, binary P matrices). As long as there is still heterogeneity in the population there is selection for better adapted and more efficient languages. Once each individual speaks the same language, there is no longer the possibility of change. Errors, however, introduce a constant degree of variation upon which natural selection for more efficient languages can act. Therefore, errors in language learning can lead to the evolution of a more efficient language. There is an interesting dependency on the error rate: small amounts of noise do not provide enough variability, large amounts of noise introduce too much variability, in between there is an optimum error rate.

The benefit of noise for language evolution is reminiscent of other situations where noise can improve the optimization properties of a system: in many evolutionary models there are optimum, non-zero mutation rates; optimum levels of noise can jiggle neural networks out of local optima; noise is the basis of the optimization method called simulated annealing.

In contrast, errors have only detrimental effects on the language acquisition mechanism for imitating randomly chosen individuals. Figure 5 shows computer simulations which repeat the simulations of Fig. 3(a) with noise. Individuals learn their language from one randomly chosen individual from the previous generation ($K = 1$). Four levels of noise are studied. For $\rho = 10^{-5}$ the evolution to a common language is essentially unaffected. For $\rho = 10^{-4}$ the system converges to a common language, but there are occasional fluctuations that generally reduce the average payoff of the population. For $\rho = 10^{-3}$ convergence to a common language is still observed, but variation and fluctuations in the average payoff are there all the time and dominate the scenario. For $\rho = 10^{-2}$ (or greater) there is no longer any convergence to a common language. Remember that for this level of error, $\rho = 10^{-2}$, the system

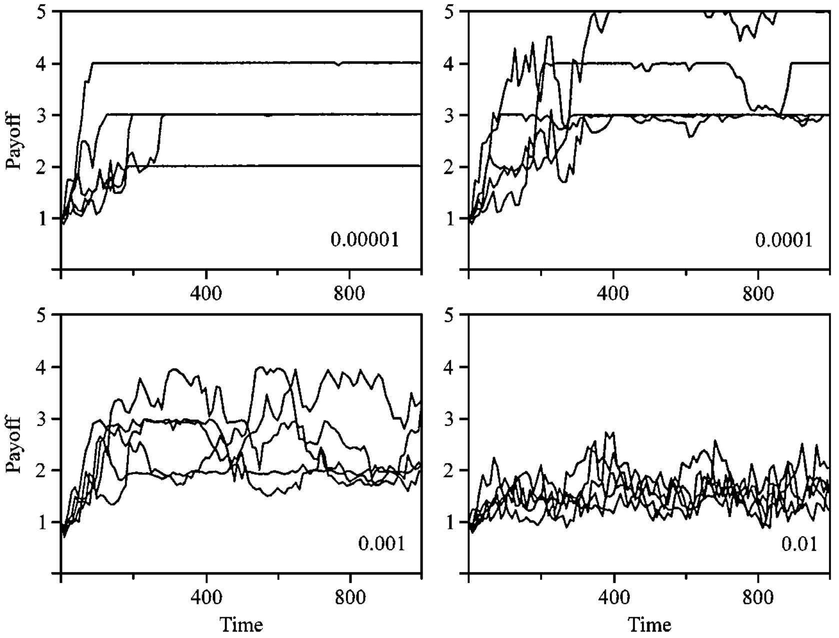

FIG. 5. Increasing amounts of noise make random sampling ineffective. Here the simulations of Fig. 3 are repeated for *rho* $= 0.00001, 0.0001, 0.001$ and 0.01. For small values of ρ the system manages to converge to a common language (which is represented a randomly chosen binary active matrix). For larger values of ρ the population no longer converges to a common language. Thus errors during language acquisition have very different effects on systems where individuals sample their parents (or other individuals with high payoff) as opposed to systems where individuals sample randomly chosen individuals. Five runs are shown for each value of ρ. For all runs we have $K = 1$.

where individuals imitate their parents always converged to optimum fitness (see Fig. 4).

Therefore, the quasi-accidental evolution of a common language which emerges when individuals imitate other randomly chosen individuals is impaired by mistakes during language acquisition. An interesting task here would be to develop an analytic framework for understanding the error-threshold phenomenon that operates here: below a critical level of ρ the evolution of a common language is possible; above this level coherence is no longer achieved.

6. Competition among Different Strategies

We have considered the performance of different mechanisms of language acquisition and described how efficiently they can drive the evolution of a common language. Now, we analyse the competition among individuals using different language acquisition devices.

Figure 6 tests parental vs. random learning. "Random players" (that learn the language from randomly chosen individuals of the previous generation) compete with players that learn the language of their parents. The error rate is 0. The payoff in the language game is interpreted as fitness. Individuals with a higher payoff leave more offspring which inherit the language acquisition strategy of their parents. Initially, 80% of individuals are random players; 20% learn the language of their parents. Initially, everyone starts off with random P matrices. As a common language evolves (that is as the average payoff increases) the fraction of random players declines. Thus individuals who learn their parents' language seem to acquire an efficient language more rapidly. This is not unexpected, because these individuals learn the language of someone who was at least fit enough to reproduce. They have a higher chance of acquiring a better adapted language. Once everyone speaks the same language, however, there is no longer selection against the random players. Neutral drift determines the final outcome.

Figure 7 repeats the simulation of Fig. 6 with an error rate of $\rho = 0.001$. In all 10 runs the random players become extinct. Furthermore, in

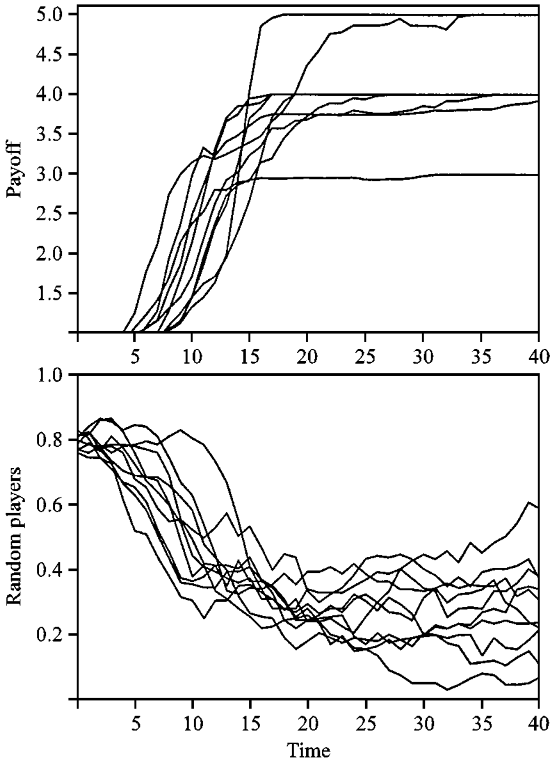

FIG. 6. It is better to learn from your parents than from random members of the population. The simulation compares the performance of two different strategies: (i) individuals who learn the language of their parents; and (ii) individuals who learn the language of an arbitrary member of the community (so called "random players"). The population size is 200. Initially 80% of the population are random players. The figure shows the increase in the average payoff as the community evolves a common language and the decrease of the fraction of random players. Note that the frequency of random players decline as long as the population is in the state of improving its language; once everyone speaks the same language there is no longer a difference in the performance of the two strategies (and random drift will eventually drive one of them to extinction). Parameters: $n = 5$, $m = 5$, every individual samples one other individual once (i.e. $K = 1$ and $k = 1$), 10 runs are shown.

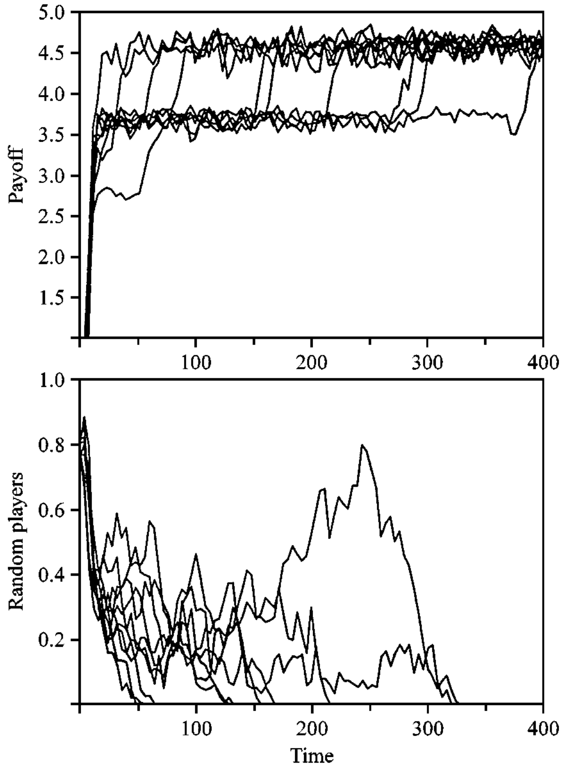

FIG. 7. Take advantage of mistakes. The possibility of errors during language acquisition enables the system to reach the global optimum. Furthermore, such errors introduce a consistent selection pressure against the random players. The figure shows 10 simulation for exactly the same parameter values as in Fig. 4. The only difference is that individuals can make mistakes while learning the language: with probability ρ, the response of another individual to a particular object is mistaken for another (randomly chosen) response, We use $\rho = 0.01$. In 10 of 10 runs, the random players become extinct and the population converges to the maximum payoff were each object is associated with exactly one signal and vice versa.

all 10 cases the population converges to the maximum fitness, where each object is associated with exactly one word. Errors make a difference.

In Fig. 8, we expand this theme and analyse competition among strategies of language acquisition that are described by integers M, such that an individual playing strategy M learns its language by sampling its parent together with M randomly chosen individuals. We consider the strategies $M = 0, 1, 2, 3$ and 4. $M = 0$ denotes the strategy where individuals only learn from their parents. Starting with equal proportions of all strategies, we find that $M = 0$ wins in all 10 runs. All other strategies become extinct. Thus learning from *randomly* chosen individuals in addition to your parents seems to be disadvantegeous.

Figures 6–8 can also be interpreted as analysing the competition between role model and random learning. (Remember that for $K = 1$ the algorithms for biological and cultural transmission are identical.) Thus, learning the language from individuals with a high payoff in the language game is better than learning from random individuals. Hence, both role model and parental learning outperform random learning.

But what about parental vs. role model learning? Clearly, there is no difference in our model

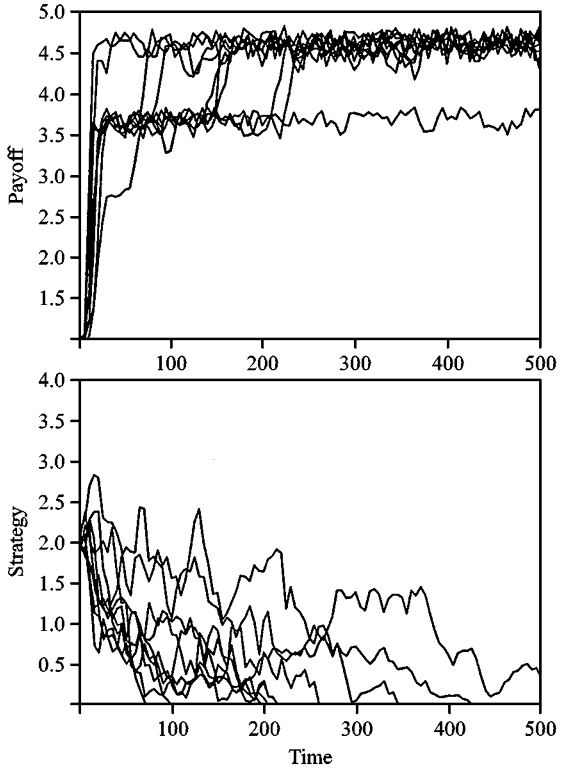

FIG. 8. In the presence of errors, is it better only to acquire the language from the parents or acquire the language from the parents and some other, randomly chosen individuals? Strategy, M, denotes the number of individuals that are sampled in addition to the parents. That is, every individual samples its parents and M other individuals of the previous generation. In our simulation, M varies from 0 to 4. In 10 out of 10 runs the average strategy of the population eventually converges to 0. In all runs the average payoff eventually converges to its maximum value. Same parameters as in Fig. 5.

between learning from the parent or one other individual chosen proportional to its payoff, but is it better to imitate only the parent or the parent and some other players proportional to their payoff? The question is equivalent to whether it is better to have only one or several role models. In Section 4.2 (and Fig. 2), we saw that learning from one role model led to a faster optimization, but learning from more than one role model usually led to a higher average fitness. Here we study the direct competition among these strategies and include errors. We performed the same simulations as for Fig. 8, analysing the competition among strategies that learn from their parents and M other individuals of the previous generation which are selected proportional to their payoff. Again using the concept of errors during language acquisition (with $\rho = 0.001$), we performed 100 runs starting with equal proportion of all strategies, M ranging from 0 to 4. The outcome is inconclusive. Fifty-five runs converged to a homogeneous population with $M = 0$, whereas 36 runs converged to homogeneous populations with M either 1, 2, 3 or 4. (Nine runs still contained a mixture of strategies after 1000 generations.) Hence, it does not seem to make much difference whether the language is learned from one relatively successful individual (the parent) or several. A clearer understanding of the differential effects of such strategies means considering repeated rounds of learning and communication during each generation.

An interesting issue in linguistics is the difference between learning the language from the parents vs. learning from other individuals. While early language acquisition is clearly from the parents, later on children seem to adopt the language and accent of their peers, not their parents. A possible explanation is that language serves as an important tool of identification between and within peer groups. While it should be possible to explore this fascinating question with the mathematical framework developed here, it is beyond the scope of the current paper.

7. Conclusions

We presented a general mathematical framework for studying the evolution of a simple communication system. Starting from random initial conditions we demonstrated how a population can evolve a common language, where specific words are associated with specific objects. The words that evolve in our simulations are arbitrary, discrete and referential.

We analysed the evolutionary properties of three mechanism of language acquisition: (i) Parental learning implies that children learn the language of their parents. (ii) Role model learning is based on the concept that the payoff in the language game is related to reputation. Children learn the language of individuals with a high reputation. (iii) Random learning describes the situation where children learn the language from randomly chosen individuals from the previous generation.

Only for parental learning do we directly assume that language confers increased biological fitness (better communicators have more offspring). For role model learning, language confers reputation (better communicators have more immitators), while for random transmission the payoff in the language game appears to be totally irrelevant. Our results demonstrate that parental learning and role model learning are effectively indistinguishable, and are in fact identical when individuals learn from a single role model. (However, our model does not consider deceit, and this will be minimized by either kin selection acting on parental learning, or in structured populations, where limited mobility favours honesty.)

Random learning consistently underperforms both parental learning and role model learning. In fact, random transmission can be modelled without any payoff evaluation; the individuals only have to imitate others. It is important, however, to emphasize that for the evolutionary stability of language there must always be an implicit assumption that language confers biological fitness. Otherwise, individuals who opt not to participate in language learning and communication, would be neutral and therefore would not be selected against. This in effect states that individuals who speak a sort of gibberish, could eventually come to dominate the population. While it is very difficult to relate the complexities of language to biological fitness, it is not so difficult to find examples where linguistic imperfections exclude speakers from dominant (cultural) positions. It becomes a separate issue to relate status to long-term reproductive success (in humans).

Without errors during language acquisition, all three mechanisms can lead to a common language. The final state is reached when each object is associated with exactly one signal. The same signal, however, may be associated with several objects. Thus our models provide one simple account for why true synonyms should be rare, while homonyms are plentiful. This discrepancy arises out of the dynamics of the learning process, in particular, the fact that synonyms are expected to disappear over several learning trials.

For parental and role model learning, the final state is optimized for performance; in the process of convergence toward the final state there is selection for languages with higher payoff. For random learning, this is not the case, and the final state is a completely randomly chosen binary matrix. There is, however, some accidental optimization for random learning, because a randomly chosen binary P matrix leads on an average to higher payoff than a random P matrix whose entries are real numbers between 0 and 1. (Hurford, 1989, used random transmission to study the acquisition of signs and hence his simulations are to a large extent based on this optimization property of the stochastic process.)

By allowing errors during language acquisition there is strictly speaking no final state and evolution will continue indefinitely. For parental and role model learning, there is an optimum error rate that maximizes the performance of the system: the chance of reaching a global optimum is increased as the system does not remain at local optima. This is of course analogous to an annealing process, whereby stochasticity can prevent fixation on local optima. The interesting question is whether children employ low levels of error, in an effort to make learning more efficient. Thus mistakes during acquisition would be viewed as an essential, rather than incidental component of language learning. In contrast, random learning is impaired by the presence of errors: if the error rate is above a small threshold value random learning does not work at all.

The interesting complications, which lead to different possibilities for formulating a language game dynamics, arise here because the language of an individual (P and Q), determines its fitness, but what is inherited from the parents is a mechanism for acquiring a language—a language acquisition device. In standard evolutionary game theory (Maynard Smith, 1982; Sigmund, 1995; Hofbauer & Sigmund, 1998), offspring inherit a strategy from their parents and this strategy determines the individual's (frequency-dependent) fitness. In the language game, children inherit a device for acquiring a language from others in the population. Thus, the language that they learn depends on what others are doing and is therefore frequency dependent. In the evolutionary language game, not only the fitness of individuals, but also their language is frequency dependent. Thus there is the possibility for decoupling cultural and biological fitness. This could lead to the emergence of specific linguistic practises that were biologically deleterious, nevertheless maintained by biological selection on the acquisition strategy.

We thank the two anonymous referees for their constructive suggestions. Support from The Florence Gould Foundation, J. Seward Johnson, Sr. Charitable Trusts, The Ambrose Monell Foundation and The Alfred P. Sloan Foundation is gratefully acknowledged.

REFERENCES

Anisfield, M. (1984). *Language Development from Birth to Three*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Aoki, K. & Feldman, M. W. (1987). *Proc. Natl. Acad. Sci.* **84,** 7164–7168.

Aoki, K. & Feldman, M. W. (1989). *Theor. Popu. Biol.* **35,** 181–194.

Balaban, E. (1988). Bird song syntax: learned intraspecific variation is meaningful. *Proc. Roy. Soc. Lond. B* **85,** 3657–3660.

Bates, E. (1992). Language development. *Current Opinions Neurobiol.* **2,** 180–185.

Bickerton, D. (1990). *Language and Species*. Chicago: University of Chicago Press.

Boyd, R. & Richardson, P. J. (1985). *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.

Bradbury, J. W. & Vehrencamp, S. L. (1998). *Principles of Animal Communication*. MA: Sinauer.

Briscoe, E. J. (1999) Grammatical acquisitions. *Language*. (In press).

Brown, R. (1973). *A First Language*. Cambridge, MA: Harvard University Press.

Burling, R. (1989). Primate calls, human language, and nonverbal communication. *Current Anthropol.* **34,** 25–53.

Cavalli-Sforza, L. L. & Feldman, M. W. (1981). *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton, NJ: Princeton University Press.

Cavalli-Sforza, L. L. & Cavalli-Sforza, F. (1995). *The Great Human Diasporas*. Reading, MA: Addison-Wesley.

Cavalli-Sforza, L. L. (1997). *Proc. Natl. Acad. Sci.* **94,** 7719–7724.

Catchpole, C. K., Leisler, B. & Dittami, J. (1984). Differential responses to male song repertoires in female songbirds implanted with oestradiol. *Nature* **312,** 563–564.

Cheney, D. L. & Seyfarth, R. M. (1990). *How Monkeys see the World: Inside the Mind of Another Species*. Chicago: Chicago University Press.

Chomsky, N. (1965). *Aspects of the Theories of Syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (1972). *Language and Mind*. New York: Harcourt Brace.

Chomsky, N. (1975). *Reflections on Language*. New York: Pantheon Press.

Chomsky, N. (1980). *Rules and Representations*. New York: Columbia University Press.

Deacon, T. (1997). The Symbolic Species. London: Allen Lane, The Penguin Press.

Dunbar, R. (1997). *Grooming, Gossip and the Evolution of Language*. Cambridge, MA: Harvard University Press.

Frisch, K. (1967). *The Dance Language and Orientation of Bees*. Cambridge, MA: Harvard University Press.

Galef, B. G. J. (1988). Imitation in animals: history, definitions and interpretation of data from the psychological laboratory. In: *Social Learning: Psychological and Biological Perspectives* (Zentall, T. & Galef, B. G., eds), pp. 3–28. Hillsdale, NJ: Lawrence Erlbaum Associates.

Greenberg, J. H. (1971). *Language, Culture and Communication*. Stanford, CA: Stanford University Press.

Greenfield, P. (1991). Imitation, grammatical development, and the invention or protogramar by an ape. In: *Biobehavioral Foundations of Language Development*. (Krasnegor, N., Rumbaugh, D. M., Studdert-Kennedy, M. & Scheifelbusch, D. eds). Berlin: Springer.

Hashimoto, T. & Ikegami, T. (1995). Lecture Notes in Artificial Intelligence. Vol. 929, pp. 812–823.

Hauser, M. D. (1996). *The Evolution of Communication*. Cambridge, MA: Harvard University Press.

Hofbauer, J. & Sigmund, K. (1998). *Evolutionary Games and Replicator Dynamics*. Cambridge: Cambridge University Press.

Hutchins, E. & Hazelhurst, B. (1995). *How to Invent a Lexicon: The Development of Shared Symbols in Interaction*. London: UCL.

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

Hurford, J. R. (1991). The evolution of the critical period for language acquisition. *Cognition* **40,** 159–201.

Hurford, J. R., Studdert-Kennedy, M. & Knight, C. (1988). *Approaches to the Evolution of Language*. Cambridge: Cambridge University Press.

Kroodsma, D. E. & Konishi, M. (1991). A suboscine bird (eastern phoebe, sayornis phoebe) develops normal song without auditory feedback. *Animal Behav.* **42,** 477–487.

Lemon, R. E. (1975). How birds develop song dialects. *Condor*. **77,** 385–406.

Lieberman, P. (1984). *The Biology and Evolution of Language*. Cambridge, MA: Harvard University Press.

Lieberman, P. (1991). *Uniquely Human*. Cambridge, MA: Harvard University Press.

MacLennan, B. (1992). In: *Artificial Life II: SFI Studies in the Sciences of Complexity* (Langton, C. G., Taylor, C. D. F. & Rasmussen, S., eds), pp. 631–658. Redwood City, CA: Addison-Wesley.

Marler, P. (1970). Birdsong and speech development: could there be parallels? *Amer. Scientist* **58,** 669–673.

Marler, P. & Pickert, R. (1984). Species-universal microstructure in the learned song of the swamp sparrow (*Melospiza georgiana*). *Animal Behav.* **32,** 673–689.

Marler, P., Peters, S., Ball, G. F., Dufty, A. M. & Wingfield, J. C. (1988). The role of sex steriods in the acquisition and production of birdsong. *Nature* **336,** 770–772.

Maynard Smith, J. (1982). *Evolution and the Theory of Games*. Cambridge: Cambridge University Press.

Maynard Smith, J. & Szathmary, E. (1995). *The Major Transitions in Evolution*. New York: Freeman.

Meltzhoff, A. N. (1993). The centrality of motor coordination and proprioception in social and cognitive development: from shared actions to shared minds. In: *The Development of Coordination in Infancy*. (Savelsbergh, G. J. P., ed.), pp. 463–496, Amsterdam: North-Holland.

MELTZHOFF, A. N. & MOORE, M. K. (1994). Imitation, memory, and the representation of persons. *Infant Behav. Dev.* **17,** 83–100.

NEWPORT, E. L. & MEIER, R. P. (1985). The acquisition of American Sign Language. In: *The Crosslinguistic Study of Language Acquisition.* Vol. 1, *The Data.* (Slobin, D. I., ed.), pp. 881–938. Hillsdale, NJ: Lawrence Erlbaum Associates.

NEWPORT, E. (1990). Maturational constraints on language learning. *Cognitive Sci.* **14,** 11–28.

NOBRE, A., ALLISON, T. & MCCARTHY, G. (1994). Word recognition in the human inferior temporal lobe. *Nature* **372,** 260–263.

NOWAK, M. A. & KRAKAUER, D. C. (1999). The evolution of Language. *Proc. Natl. Acad. Sci. U.S.A.* **96,** 8028–8033.

OLIPHANT, M. (1996). The dilemma of Saussurean communication. *BioSystems* **37,** 31–38.

PALACIOS, O. A., STEPHENS, C. R. & WAELBROECK, H. (1998). Emergence of algorithmic language in genetic systems. *Biosystems* **47,** 129–147.

PARISI, D. (1997). Artificial life and higher level cognition. *Brain Cognition* **34,** 160–184.

PINKER, S. & BLOOM, P. (1990). Natural language and natural selection. *Behavioural Brain Sci.* **13,** 707–784.

PINKER, S. (1994). *The Language Instinct.* New York: Marrow.

PORAN, N. S. & COSS, R. G. (1990). Development of antisnake defenses in California ground squirrels (*Spermophilus beecheyi*): I. Behavioural and immunological relationships. *Behaviour* **112,** 222–245.

SACHS, J. (1985). Prelinguistic development. In: *The Development of Language* (Gleason, J. B. ed.). OH: Charles E. Merrill Publishing Co. Columbus.

SAUSSURE, F. de. (1916). *Cours de Linguistique Generale.* Paris: Paycot.

SIGMUND, K. (1995). *Games of Life.* Harmondsworth: Penguin.

SEYFARTH, R., CHENEY, D. & MARLER, P. (1980). Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science* **210,** 801–803.

SMITH, W. J. (1977). *The Behaviour of Communicating.* Cambridge, MA: Harvard University Press.

STEELS, L. (1997). *The Synthetic Modelling of Language Origins. Evolution of Communication Journal.* Vol: 1(1). Amsterdam: John Benjamins.

SZATHMARY, E. & MAYNARD SMITH, J. (1995). The major evolutionary transitions. *Nature* **374,** 227–232.

WITTELSON, S. & PALLIE, W. (1973). Left hemispheric specialization for language in the newborn: neuroanatomical evidence of assymetry. *Brain* **96,** 641–646.

APPENDIX A

Counting Random Binary Matrices

In this appendix, we describe how to count and classify all binary P matrices, which are the absorbing states for the dynamical systems discussed in Section 4. In other words, we focus our attention on $n \times m$ binary matrices with a single 1 in each row. There are clearly m^n such matrices. The more difficult combinatorics of the classification occur in the case when $n \leqslant m$, which we will now consider.

We would like to classify these matrices by the fitness which they yield. Let us denote the number of matrices yielding fitness i by Z_i. The maximum possible fitness is, of course, $2n$. There are clearly $Z_n = m!/(m-n)!$ such maximal-fitness matrices. However, we also want expressions for the number of sub-optimal matrices, i.e. for Z_{n-1}, $Z_{n-2}, \ldots, Z_1$. Notice that if we count each Z_{n-k} correctly then we will have the equality $\sum_{k=0}^{k=n-1} Z_{n-k} = m^n$.

Working from the definition of Z_{n-k}, we derive the following expression for the number of matrices yielding fitness $n-k$:

$$Z_{n-k} = \sum_{t_1, t_2, \ldots, t_m} \binom{n}{t_1+1} \binom{n-(t_1+1)}{t_2+1} \cdots \binom{n - \sum_{i=1}^{m-1}(t_i+1)}{t_m+1} \frac{(m-a)!}{(m-n+k)!},$$

where we sum over all partitions $t_1, \ldots, t_m$ satisfying:

$$t_i \geqslant 0,$$

$$t_1 + t_2 + \cdots t_m = k,$$

$$t_i \leqslant n-1,$$

$$k + a \leqslant n,$$

$$\text{where } a = \#\{t_i | t_i > 0\}.$$

Moreover, in each term of the expression above we only want to include the factors involving t_i and $t_i > 0$. We can rewrite our expression for Z_{n-k} as

$$Z_{n-k} = \frac{n!}{(m-n+k)!} \times \sum_{\{t_1, \ldots, t_m\}} \frac{(m-a)!}{(t_1+1)!(t_2+1)! \cdots (t_m+1)!(n-k-a)!},$$

where the sum is taken over the same set of partitions t_i as before.

Table 1 shows numerical results for $n = 4$ and $m = 4, 5, \ldots, 15$. For each value of m we report the average fitness,

$$\bar{f} = \frac{2}{m^n} \sum_{i=1}^{n} i \cdot Z_i,$$

as well as the proportion of matrices which have maximum fitness,

$$P_{n,m} = \frac{Z_n}{m^n}.$$

We see that the average fitness, $\bar{f}$, approaches the maximum value $2n$ for increasing values of m. There is a simple, analytic explanation for this trend. When n is fixed, notice that

$$\begin{aligned}\lim_{m\to\infty} P_{n,m} &= \lim_{m\to\infty} \frac{Z_n}{m^n} \\ &= \lim_{m\to\infty} \frac{m!}{(m-n)!(m^n)} \\ &= \lim_{m\to\infty} \frac{(m)(m-1)(m-2)\cdots(m-n+1)}{m^n} \\ &= \lim_{m\to\infty} \frac{m^n + \text{lower powers of } m}{m^n} \\ &= 1.\end{aligned}$$

In other words, as m gets large, the fraction of matrices which yield maximum fitness approaches 1. As a result, the average fitness approaches the maximum fitness, n. Thus, learning the language from random individuals irrespective of their payoff in the language game can lead to a fairly efficient language especially if m is significantly larger than n.

As a specific example, we consider $n = m = 4$. In this case,

$$Z_n = 4! = 24.$$

For Z_{n-1} we consider the four partitions

$$1 = 1 + 0 + 0 + 0,$$

$$1 = 0 + 1 + 0 + 0, \text{ etc.}$$

to obtain

$$Z_{n-1} = 4 \cdot \left(\frac{4!}{1!}\frac{3!}{2!2!}\right) = 144.$$

For Z_{n-2}, we must consider the four partitions

$$2 = 2 + 0 + 0 + 0,$$

$$2 = 0 + 2 + 0 + 0, \text{ etc.}$$

and also the $\binom{4}{2}$ partitions

$$2 = 1 + 1 + 0 + 0,$$

$$2 = 1 + 0 + 1 + 0, \text{ etc.}$$

to obtain

$$Z_{n-2} = \frac{4!}{2!}\left(4 \cdot \frac{3!}{3!1!} + \binom{4}{2}\frac{2!}{2!2!0!}\right) = 84.$$

For Z_{n-3}, we consider all the partitions

$$3 = 3 + 0 + 0 + 0, \text{ etc.}$$

$$3 = 2 + 1 + 0 + 0, \text{ etc.}$$

$$3 = 1 + 1 + 1 + 0, \text{ etc.}$$

But notice that only the partitions of 3 in which $a = 1$ satisfy $k + a \leqslant n$, so that according to our formula

$$Z_{n-3} = 4 \cdot \left(\frac{4!}{3!}\frac{3!}{4!0!}\right) = 4.$$

Finally, we check that $Z_n + Z_{n-1} + Z_{n-2} + Z_{n-3} = 256 = m^n$, as desired.

APPENDIX B

A Paradoxical Result on Mutual Understanding

If P and Q are unrelated in an individual, then it is easily possible to construct 2 languages L_1 and L_2 that understand each other better than themselves. Thus, we may have

$$F(L_1, L_1) < F(L_1, L_2) = F(L_2, L_1) > F(L_2, L_2). \tag{7}$$

A simple example is

$$P_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad Q_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

and

$$P_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Q_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Here, $F(L_1, L_1) = F(L_2, L_2) = 0$ whereas $F(L_1, L_2) = F(L_2, L_1) = 2$.

An interesting question is how closely related P and Q must be such that this paradoxical situation is not possible?

Note that the coupling of P and Q via an association matrix, see eqn (9), is not strong enough to exclude this possibility. An example is given by the 2 association matrices

$$A_1 = \begin{pmatrix} 23 & 77 \\ 66 & 34 \end{pmatrix} \quad \text{and} \quad A_2 = \begin{pmatrix} 1 & 99 \\ 34 & 66 \end{pmatrix}.$$

If P_1, Q_1 and P_2, Q_2 are the constructed according to eqn (9) then we find $F(L_1, L_1) = 2.3743..$, $F(L_1, L_2) = F(L_2, L_1) = 2.3894..$; and $F(L_2, L_2) = 2.3771...$ Hence L_1 and L_2 understand each other better than themselves. In evolutionary game dynamics the two languages are in equilibrium. Both languages have an advantage when rare!