

# Spontaneous Lexicon Change

Luc Steels<sup>(1,2)</sup> and Frédéric Kaplan<sup>(1,3)</sup>

(1) Sony CSL Paris - 6 Rue Amyot, 75005 Paris

(2) VUB AI Lab - Brussels

(3) LIP6 - 4, place Jussieu 75232 Paris cedex 05

## Abstract

The paper argues that language change can be explained through the stochasticity observed in real-world natural language use. This thesis is demonstrated by modeling language use through language games played in an evolving population of agents. We show that the artificial languages which the agents spontaneously develop based on self-organisation, do not evolve even if the population is changing. Then we introduce stochasticity in language use and show that this leads to a constant innovation (new forms and new form-meaning associations) and a maintenance of variation in the population, if the agents are tolerant to variation. Some of these variations overtake existing linguistic conventions, particularly in changing populations, thus explaining lexicon change.

## 1 Introduction

Natural language evolution takes place at all levels of language (McMahon, 1994). This is partly due to external factors such as language contact between different populations or the need to express new meanings or support new modes of interaction with language. But it is well-established that language also changes spontaneously based on an internal dynamics (Labov, 1994). For example, many sound changes, like from /b/ to /p/, /d/ to /t/, and /g/ to /k/, which took place in the evolution from proto-Indo-European to Modern Germanic languages, do not have an external motivation. Neither do many shifts in the expression of meanings. For example, the expression of future tense in English has shifted from "shall" to "will", even though "shall" was perfectly suited and "will" meant something else

(namely "wanting to"). Similarly, restructuring of the grammar occurs without any apparent reason. For example, in Modern English the auxiliaries come before the main verb, whereas in Old English after it ('he conquered be would' (Old English) vs. 'he would be conquered' (Modern English)). This internal, apparently non-functional evolution of language has been discussed widely in the linguistic literature, leading some linguists to strongly reject the possibility of evolutionary explanations of language (Chomsky, 1990).

In biological systems, evolution takes place because [1] a population shows natural variation, and [2] the distribution of traits in the population changes under the influence of selection pressures present in the environment. Note that biological variation is also non-functional. Natural selection acts *post factum* as a selecting agent, pushing the population in certain directions, but the novelty is created independently of a particular goal by stochastic forces operating during genetic transmission and development. Our hypothesis is that the same applies to language, not at the genetic but at the cultural level. We hypothesise that language formation and evolution take place at the level of language itself, without any change in the genetic make up of the agents. Language recruits and exploits available brain capacities of the agents but does not require any capacity which is not already needed for other activities (see also (Batali, 1998), (Kirby and Hurford, 1997)).

The present paper focuses on the lexicon. It proposes a model to explain spontaneous lexicon evolution, driven solely by internal factors. In order to have any explanatory force at

all, we cannot put into the model the ingredients that we try to explain. Innovation, maintenance of variation, and change should follow as emergent properties of the operation of the model. Obtaining variation is not obvious, because a language community should also have a natural tendency towards coherence, otherwise communication would not be effective. An adequate explanatory model of lexicon change must therefore show [1] how a coherent lexicon may arise in a group of agents, [2] how nevertheless the lexicon may remain internally varied and exhibit constant innovation, and [3] how some of this variation may be amplified to become dominant in the population. These three quite difficult challenges are taken up in the next three sections of the paper.

## 2 How a coherent lexicon may arise

To investigate concretely how a lexicon may originate, be transmitted from one generation to the next, and evolve, we have developed a minimal model of language use in a dynamically evolving population, called the *naming game* (Steels, 1996). The naming game has been explored through computational simulations and is related to systems proposed and investigated by (Oliphant, 1996), (MacLennan, 1991), (Werner and Dyer, 1991), a.o. It has even been implemented on robotic agents who develop autonomously a shared lexicon grounded in their sensori-motor experiences (Steels and Vogt, 1997), (Steels, 1997). The naming game focuses on associating form and meaning. Obviously in human natural languages both form and meaning are non-atomic entities with complex internal structure, but the results reported here do not depend on this internal complexity.

We assume a set of *agents*  $\mathcal{A}$  where each agent  $a \in \mathcal{A}$  has contact with a set of *objects*  $\mathcal{O} = \{o_0, \dots, o_n\}$ . The set of objects constitutes the environment of the agents. A *word* is a sequence of letters randomly drawn from a finite alphabet. The agents are all assumed to share the same alphabet. A *lexicon*  $\mathcal{L}$  is a time-dependent relation between objects, words, and a score. Each agent  $a \in \mathcal{A}$  has his own set of words  $W_{a,t}$  and his own lexicon

$L_{a,t} \subset \mathcal{O}_a \times W_{a,t} \times \mathcal{N}$ , which is initially empty. An agent  $a$  is therefore defined at a time  $t$  as a pair  $a_t = \langle W_{a,t}, L_{a,t} \rangle$ . There is the possibility of synonymy and homonymy: An agent can associate a single word with several objects and a given object with several words. It is not required that all agents have at all times the same set of words and the same lexicon.

We assume that environmental conditions identify a context  $C \subset \mathcal{O}$ . The speaker selects one object as the topic of this context  $f_s \in C$ . He signals this topic using extra-linguistic communication (such as through pointing). Based on the interpretation of this signalling, the hearer constructs an object score  $0.0 \leq e_o \leq 1.0$  for each object  $o \in C$  reflecting the likelihood that  $o$  is the speaker's topic. If there is absolute certainty, one object has a score of 1.0 and the others are all 0.0. If there is no extra-linguistic communication, the likelihood of all objects is the same. If there is only vague extra-linguistic communication, the hearer has some idea what the topic is, but with less certainty. The *meaning scope* parameter  $\sigma_m$  determines the number of object candidates the hearer is willing to consider. The *meaning focus* parameter  $\phi_m$  determines the tolerance to consider objects that are not the center of where the speaker pointed to. In the experiments reported in this paper, the object-score is determined by assuming that all objects are positioned on a 2-dimensional grid. The distance  $d$  between the topic and the other objects determines the object-score, such that

$$e_{object} = \frac{1}{1 + \left(\frac{d}{\phi_m}\right)^2} \quad (1)$$

$\phi_m$  is the *meaning focus* factor.

To name the topic, the speaker retrieves from his lexicon all the associations which involve  $f_s$ . This set is called the association-set of  $f_s$ . Let  $o \in \mathcal{O}$  be an object,  $a \in \mathcal{A}$  be an agent, and  $t$  a time moment, then the association-set of  $o$  is

$$A_{o,a,t} = \{ \langle o, w, u \rangle \mid \langle o, w, u \rangle \in L_{a,t} \} \quad (2)$$

Each of the associations in this set suggests a word  $w$  to use for identifying  $o$  with a score

$0.0 \leq u \leq 1.0$ . The speaker orders the words based on these scores. He then chooses the association with the largest score and transmits the word which is part of this association to the hearer.

Next the hearer receives the word  $w$  transmitted by the speaker. To handle stochasticity the hearer not only considers the word itself a set of candidate words  $W$  related to  $w$ . These are all the words in the word-set of the hearer  $W_{h,t}$  that are either equal to  $w$  or related with some distance to  $w$ . The *form scope* parameter  $\sigma_f$  determines how far this distance can be. A score is imposed over the members of the set of candidate words:

$$m_{w_1} = \frac{1}{1 + \left(\frac{d}{\phi_f}\right)^2} \quad (3)$$

$\phi_f$  is the form-focus factor. The higher this factor, the sharper the hearer has been able to identify the word produced by the speaker, and therefore the less tolerant the hearer is going to be to accept other candidates.

For each word  $w_j$  in  $W$ , the hearer then retrieves the association-set that contains it. He constructs a *score-matrix* which contains for each object a row and for each word-form a column. The first column contains the object-scores  $e_o$ , the first row the form-scores  $m_{w_j}$ . Each cell in the inner-matrix contains the association-score for the relation between the object and the word-form in the lexicon of the hearer:

		$w_1$	$w_2$	...
		$m_{w_1}$	$m_{w_2}$	...
$o_1$	$e_{o_1}$	$u_{\langle o_1, w_1 \rangle}$	$u_{\langle o_1, w_2 \rangle}$	...
$o_2$	$e_{o_2}$	$u_{\langle o_2, w_1 \rangle}$	$u_{\langle o_2, w_2 \rangle}$	...
...	...	...	...	...

Obviously many cells in the matrix may be empty (and then set to 0.0), because a certain relation between an object and a word-form may not be in the lexicon of the hearer. Note also that there may be objects identified by lexicon lookup which are not in the initial context  $C$ . They are added to the matrix, but their object-score is 0.0.

The final state of an inner matrix cell of the score matrix is computed by taking a weighted

sum of (1) the object-score  $e_o$  on its row, (2) the word-form score  $m_w$  on its column, and (3) the association-score  $u_{\langle o, w \rangle}$  in the cell itself. Weights indicate how strong the agent is willing to rely on each source of information. One object-word pair will have the best score and the corresponding object is the topic  $f_h$  chosen by the hearer. The association in the lexicon of this object-word pair is called the winning association. This choice integrates extra-linguistic information (the object-score), word-form ambiguity (the word-form-score), and the current state of the hearer's lexicon (the association-score).

The hearer then indicates to the speaker what topic he identified. In real-world language games, this could be through a subsequent action or through another linguistic interaction. When a decision could be made and  $f_h = f_s$  the game succeeds, otherwise it fails.

The following adaptations take place by the speaker and the hearer based on the outcome of the game.

**1. The game succeeds** This means that speaker and hearer agree on the topic. To reinforce the lexicon, the speaker increments the score  $u$  of the association that he preferred, and hence used, with a fixed quantity  $\delta$ . And decrements the score of the  $n$  competing associations with  $\delta$ . 0.0 and 1.0 remain the lower and upper-bound of  $u$ . An association is competing if it associates the topic  $f_s$  with another word. The hearer increments by  $\delta$  the score of the association that came out with the best score in the score-matrix, and decrements the  $n$  competing associations with  $\delta$ . An association is competing if it associates the wordform of the winning association with another meaning.

**2. The game fails** There are several cases:

*1. The Speaker does not know a word*

It could be that the speaker failed to retrieve from the lexicon an association covering the topic. In that case, the game fails but the speaker may create a new word-form  $w'$  and associate this with the topic  $f_s$  in his lexicon. This happens with a word creation probability  $w_c$ .

*2. The hearer does not know the word.*

In other words there is no association in the

lexicon of the hearer involving the word-form of the winning association. In that case, the game ends in failure but the hearer may extend his lexicon with a word absorption probability  $w_a$ . He associates the word-form with the highest form-score to the object with the highest object-score.

3. *There is a mismatch between  $f_h$  and  $f_s$ .*

In this case, both speaker and hearer have to adapt their lexicons. The speaker and the hearer decrement with  $\delta$  the association that they used.

Figure 1 shows that the model achieves our first objective. It displays the results of an experiment where in phase 1 a group of 20 agents develops from scratch a shared lexicon for naming 10 objects. Average game success reaches a maximum and lexicon coherence (measured as the average spread in the population of the most dominant form-meaning pairs) is high (100 %) In the early stage there is important lexicon change as new form-meaning pairs need to be generated from scratch by the agents. Lexicon change is defined to take place when a new form-meaning pair overtakes another one in the competition for the same meaning.

Phase 2 demonstrates that the lexicon is resilient to a flux in the population. An in- and outflow of agents is introduced. A new agent coming into the population has no knowledge at all about the existing set of conventions. Success and coherence therefore dip but quickly regain as the new agents acquire the existing lexicon. High coherence is maintained as well as high average game success. Between the beginning of the flux and the end (after 30,000 language games), the population has been renewed 5 times. Despite of this, the lexicon has not changed. It is transmitted across generations without change.

### 3 How a lexicon may innovate and maintain variation

So, although this model explains the formation and transmission of a lexicon it does not explain why a lexicon might change. Once a winner-take-all situation emerges, competing forms are completely suppressed and no new innovation arises. Our hypothesis is that inno-

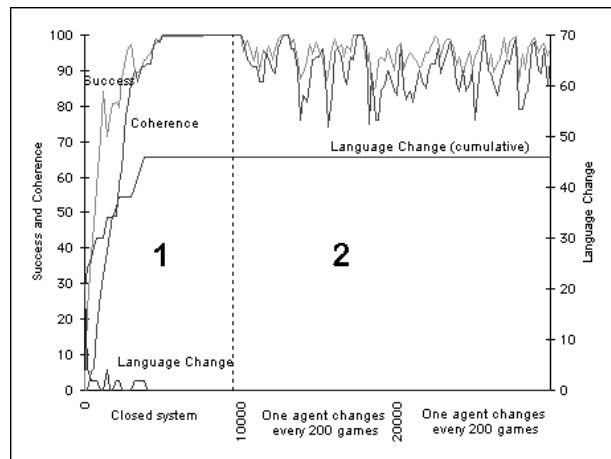


Figure 1: The graphs show for a population of 20 agents and 10 meanings how a coherent set of form-meaning pairs emerges (phase 1). In a second phase, an in- and outflow of agents (1 in/outflow per 200 games) is introduced, the language stays the same and high success and coherence is maintained.

vation and maintenance of variation is caused by stochasticity in language use (Steels and Kaplan, 1998). Stochasticity naturally arises in real world human communication and we very much experienced this in robotic experiments as well. Stochasticity is modeled by a number of additional stochastic operators:

1. Stochasticity in non-linguistic communication can be investigated by probabilistically introducing a random error as to which object is used as topic to calculate the object-score. The probability is called the topic-recognition-stochasticity  $T$ .
2. Stochasticity in the message transmission process is caused by an error in production by the speaker or an error in perception by the hearer. It is modeled with a second stochastic operator  $F$ , the form-stochasticity, which is the probability that a character in the string constituting the word form mutates.
3. Stochasticity in the lexicon is caused by errors in memory lookup by the speaker or the hearer. These errors are modeled using a third stochastic operator based on a parameter  $A$ , the memory-stochasticity,

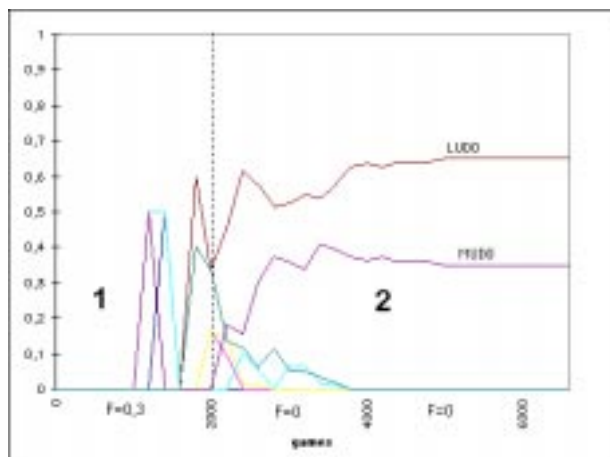


Figure 2: Competition diagram in the presence of form-stochasticity and a broad form-scope. The diagram shows all the forms competing for the same meaning and the evolution of their score. When  $F = 0.3$  new word-forms are occasionally introduced resulting in new word-meaning associations. When  $F = 0.0$  the innovation dies out although some words are still able to maintain themselves due to the hearer's broad focus.

which alters the scores of the associations in the score matrix in a probabilistic fashion.

The hearer has to take a broader scope into account in order to deal with stochasticity. He should also decrease the focus so that alternative candidates get a better chance to compete. The broader scope and the weaker focus has also the side effect that it will maintain variation in the population. This is illustrated in figure 2. In the first phase there is a high form-stochasticity as well as a broad form-scope. Different forms compete to express the same meaning and none of them manages to become the winner. When form-stochasticity is set to 0.0, the innovation dies out but the broad scope maintains both variations. One form ("ludo") emerges as the winner but another form ("mudo") is also maintained in the population. There is no longer a winner-take-all situation because agents tolerate the variation. We conclude the following:

1. Stochasticity introduces innovation in the lexicon. There is no longer a clear winner-take-

all situation, whereby the lexicon stays in an equilibrium state. Instead, there is a rich dynamics where new forms appear, new associations are established, and the domination pattern of associations is challenged. The different sources of stochasticity each innovate in their own way: Topic-stochasticity introduces new form-meaning associations for existing forms. Form-stochasticity introduces new forms and hence potentially new form-meaning associations. Memory-stochasticity shifts the balance among the word-meaning associations competing for the expression of the same meaning.

2. Tolerance to stochasticity, due to a broad scope (high  $\sigma_f$ ) and a weak focus (low  $\phi_f$ ), maintains variation. For example, suppose a form "ludo" is transmitted by the speaker but the hearer has only "mudo" in his lexicon. If the form-focus factor is low and if both forms refer in the respective agents to the same object, their communication will be successful, because the word-score of "mudo" will not deviate that much from "ludo". Neither the hearer nor the speaker will change their lexicons. Similar effects arise when the agent broadens the meaning scope and weakens its meaning focus to deal with meaning stochasticity, caused by error or uncertainty in the non-linguistic communication.

## 4 How variation is amplified

Although stochasticity and the agent's increased tolerance to cope with stochasticity explain innovation and the maintenance of variation, they do not in themselves explain lexicon change. Particularly when a language is already established, the new form-meaning pairs do not manage to overtake the dominating pair. To get lexicon change we need an additional factor that amplifies some of the variations present in the population. Several such factors are probably at work. The most obvious one is a change in the population. New agents arriving in the community may first acquire a minor variant which they then start to propagate further. After a while this variant could become in turn the dominant variant. We have conducted a series of experiments to test this hypothesis, with remarkable results. Typically there is a pe-

riod of stability (even in the presence of uncertainty and stochasticity) followed by a period of instability and strong competition, again followed by a period of stasis. This phenomenon has been observed for natural languages and is known in biology as punctuated equilibria (Eldredge and Gould, 1972).

The following are results of experiments focusing on form-stochasticity. Figure 3 shows the average game success, lexicon coherence, and lexicon change for an evolving population. 30,000 language games are shown. It starts when the population develops a lexicon from scratch (phase 1). Form-scope is constantly kept at  $\sigma_f = 5$  in other words five forms are considered similar to the world heard. Initially there is no form-stochasticity. In phase 2 a flow in the population is introduced with a new agent every 100 games. We see that there is no lexicon change. Success and Coherence is maintained at high levels. Then form-stochasticity is increased to  $\sigma_f = 0.05$  in phase 3. Initially there is still no lexicon change. But gradually the language destabilises and rapid change is observed. Interestingly enough average game success and coherence are maintained at high levels. After a certain period a new phase of stability starts.

A companion figure (figure 4) focuses on the competition between different forms for the same meaning. In the initial stage there is a winner-take-all situation (the word "bagi"). When stochasticity is present, new forms start to emerge but they are not yet competitive. It is only when the flux in the population is positive that we see one competitor "pagi" becoming strong enough to eventually overtake "bagi". "bagi" resulted from a misunderstanding of "pagi". There is a lot of instability as other words also enter into competition, giving successive dominance of "kagi", then "kugi" and then "kugo". A winner-take-all situation arises with "kugo" and therefore a new period of stability sets in. Similar results can be seen for stochasticity in non-linguistic communication and in the lexicon.

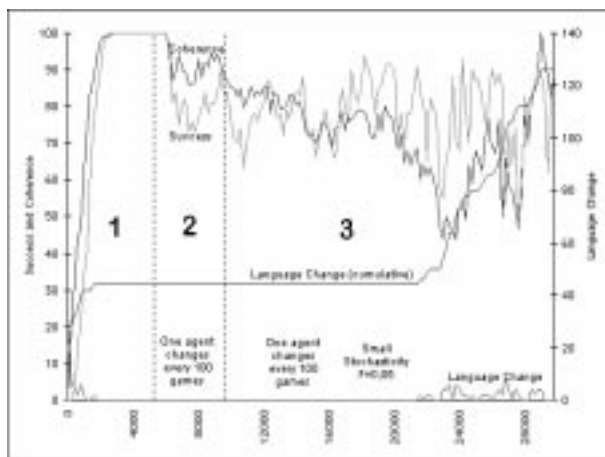


Figure 3: The diagram shows that change requires both the presence of uncertainty and stochasticity, high tolerance (due to broad scope and diffuse focus) *and* a flux in the population. The lexicon is maintained even in the case of population change (phase 2), but starts to change when stochasticity is increased (phase 3).

## 5 Conclusions

The paper has presented a theory that explains spontaneous lexicon change based on internal factors. The theory postulates that (1) coherence in language is due to self-organisation, i.e. the presence of a positive feedback loop between the choice for using a form-meaning pair and the success in using it, (2) innovation is due to stochasticity, i.e. errors in form transmission, non-linguistic communication, or memory access, (3) maintenance of variation is due to the tolerance agents need to exhibit in order to cope with stochasticity, namely the broadening of scope and the weakening of focus, and finally (4) amplification of variation happens due to change in the population. Only when all four factors are present will effective change be observed.

These hypotheses have been tested using a formal model of language use in a dynamically evolving population. The model has been implemented and subjected to extensive computational simulations, validating the hypotheses.

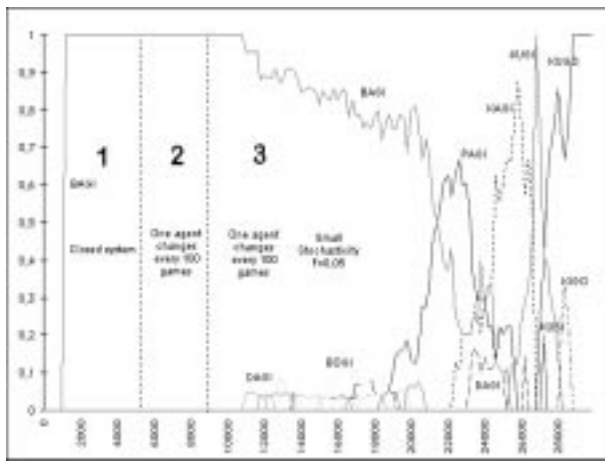


Figure 4: The diagram shows the competition between different forms for the same meaning. We clearly see first a rapid winner-take-all situation with the word "bagi", then the rise of competitors until one ("pagi") overtakes the others. A period of instability follows after which a new dominant winner ("kugo") emerges.

## 6 Acknowledgement

The research described in this paper was conducted at the Sony Computer Science Laboratory in Paris. The simulations presented have been built on top of the BABEL toolkit developed by Angus McIntyre (McIntyre, 1998) of Sony CSL. Without this superb toolkit, it would not have been possible to perform the required investigations within the time available. We are also indebted to Mario Tokoro of Sony CSL Tokyo for continuing to emphasise the importance of stochasticity in complex adaptive systems.

## References

- J. Batali. 1998. Computational simulations of the emergence of grammar. In J. Hurford, C. Knight, and M. Studdert-Kennedy, editors, *Approaches to the Evolution of Language*. Edinburgh University Press, Edinburgh.
- N. Chomsky. 1990. Rules and representations. *Brain and Behavior Science*, 3:1–15.
- N. Eldredge and S. Gould. 1972. Punctuated equilibria: an alternative to phyletic gradualism. In T. Schopf, editor, *Models in palaeobiology*, pages 82–115, San Francisco. Freeman and Cooper.
- S. Kirby and J. Hurford. 1997. Learning, culture and evolution in the origin of linguistic constraints. In P. Husbands and I. Harvey, editors, *Proceedings of the Fourth European Conference on Artificial Life*, pages 493–502. MIT Press.
- W. Labov. 1994. *Principles of Linguistic Change. Volume 1: Internal Factors*. Blackwell, Oxford.
- B. MacLennan. 1991. Synthetic ethology: An approach to the study of communication. In C. Langton, editor, *Artificial Life II*, Redwood City, Ca. Addison-Wesley Pub. Co.
- A. McIntyre. 1998. Babel: A testbed for research in origins of language. To appear in COLING-ACL 98, Montreal.
- A. McMahan. 1994. *Understanding Language Change*. Cambridge University Press, Cambridge.
- M. Oliphant. 1996. The dilemma of saurian communication. *Biosystems*, 1–2(37):31–38.
- L. Steels and F. Kaplan. 1998. Stochasticity as a source of innovation in language games. In C. Adami, R. Belew, H. Kitano, and C. Taylor, editors, *Proceedings of Artificial Life VI*, Los Angeles, June. MIT Press.
- L. Steels and P. Vogt. 1997. Grounding adaptive language games in robotic agents. In I. Harvey and P. Husbands, editors, *Proceedings of the 4th European Conference on Artificial Life*, Cambridge, MA. The MIT Press.
- L. Steels. 1996. Self-organizing vocabularies. In C. Langton, editor, *Proceeding of Alife V*, Nara, Japan.
- L. Steels. 1997. The origins of syntax in visually grounded robotic agents. In M. Pollack, editor, *Proceedings of the 15th International Joint Conference on Artificial Intelligence*, Los Angeles. Morgan Kaufmann Publishers.
- G. M. Werner and M. G. Dyer. 1991. Evolution of communication in artificial organisms. In C. G. Langton, C. Taylor, and J.D. Farmer, editors, *Artificial Life II, Vol. X of SFI Studies in the Sciences of Complexity*, Redwood City, Ca. Addison-Wesley Pub.