Semiotic schemata: Selection units for linguistic cultural evolution

Frédéric Kaplan
Sony CSL - Paris - 6 Rue Amyot, 75005 Paris
LIP6 - UPMC - 4, Place Jussieu F-75252 Paris
E-mail: kaplan@csl.sony.fr

Abstract

Words, like genes, are replicators in competition to colonize our brains. Some, by luck or thanks to their intrinsic qualities, manage to spread in entire populations. In this paper we take the approach of cultural selectionism to study the emergence of communication systems in a population of agents. By studying simple models of word competition in noisy environments, we define the basic dynamics of such systems. We then argue for their generality and introduce the notion of semiotic schemata, generic replicators that account for the different competitions that are going on during lexicon formation. Eventually, we present a synthesis of the dynamics using this new formalism.

Introduction

Genetic and cultural systems can both be seen as complex evolving dynamic architecture. In this paper we will discuss a particular paradigm for understanding the dynamics of cultural systems: *selectionism*. The comparison between genetic evolution and cultural evolution , popularized by Dawkins's *memes*, has proved to be fruitful (Dawkins, 1976; Dennett, 1995; Blackmore, 1999). By analogy with genes, Dawkins defines *memes*, as cultural replicators. Dawkins defines a replicator as "any entity in the universe which interacts with its world, including other replicators, in such way that copies of itself are made" (Dawkins, 1984). In genetics, replicators are single genes or fragments of genetic material. Evolutionary genetics study the *competition* between genetic replicators, how some of them are *selected*, how some of them *disappear*. Metaphorically, we can talk about the survival of some replicators and the death of others. From a similar perspective we could say that cultural replicators are in competition to colonize our brains. Like for genetic replicators, criteria such as *fecondity* or *fidelity* in the copying process, are useful notions to understand the victories or defeats of some memes against others.

The idea of *cultural selectionism* has been applied in the study of a particular kind of cultural evolution: the evolution of communication systems and languages. Indeed languages, like organisms, could be seen in competition with one another. They try to "survive" by being used by speakers. This particular perspective, different from contemporary linguistic approaches, is not really new. In 1937, Arsene Darmesteter wrote "the life of words" already announcing this paradigm (Darmesteter, 1937). Today, several researchers in Artificial Life talk of their work in similar terms (Batali, 1998; Hurford, 1998; Kirby, 1999a; Kirby, 1999b; Steels, 1997).

In our previous work, we have explore the emergence of complex communication systems, in particular the coupling between creation of grounded categories and lexicon formation (Steels and Kaplan, 1999c; Steels and Kaplan, 1999b; Steels and Kaplan, 1999a), and the effect of noise on the evolution of such systems (Steels and Kaplan, 1998b; Steels and Kaplan, 1998a). This paper is a synthesis, using the cultural selectionism paradigm, of the results we have obtained with complex architectures. We study basic model of linguistic evolution, simple enough to account for most experimental results in the field. In identifying these basic dynamics, we try to point out the different competitions that are going on during lexicon self-organisation. The main difficulty is to define the *right selection unit*. Like for memes, it is uneasy to precisely define what are the replicators in the cultural linguistic evolution. Should we consider competition between words, meanings or larger parts of languages?

The paper is organised as follows. In the next section, we study very simple models showing the competition between words for naming the same referent when there is noise in the environment. Then, we move to more complex architectures and analyse how these dynamics evolve when a lexicon is emerging to name a set of objects under the presence of noise during word transmission. We then argue that the dynamics identified are general and apply to other kinds of competition that are present during the self-organisation of a communication system. To account for all these different competitions, we introduce the notion of semiotic schemata, which are general replicators for linguistic evolution. We present a synthesis of the dynamics using this new formalism.

Competition between words

In this section, westudy the basic dynamics that enable one linguistic convention to be collectively chosen by a population of agents.

Positive feedback loop

Model 1.1 Each agent a in a population of N agents is defined by a single *preference vector* $(x_a^1, x_a^2, .., x_a^N)$. x_a^i represents the score that agent a gives to the word convention i. Agents interact through a very simple protocol. Two agents are picked at random in the population. One agent is speaker and the other one is hearer. When an agent is speaker, it uses the convention associated to the highest score in its *preference vector*. The convention is transmited to the hearer, and the latter simply increases by 1 the score of the convention used by the speaker in its own *preference vector*. Initially the population starts with N agents, each agent a having a single bias for one preferred convention which is modelised by a vector of size N $(0, 0, .., 0, 1, 0, .., 0)$.

Exp 1.1 (N=50, 1 run). Figure 1 shows the competition of the different word conventions for 50 agents trying to impose their word. The positive feedback loop introduced in this simple model creates a winner-take-all situation where one word dominates. The word that finally wins has no special properties and any new run of the simulation would lead to the selection of a different word.

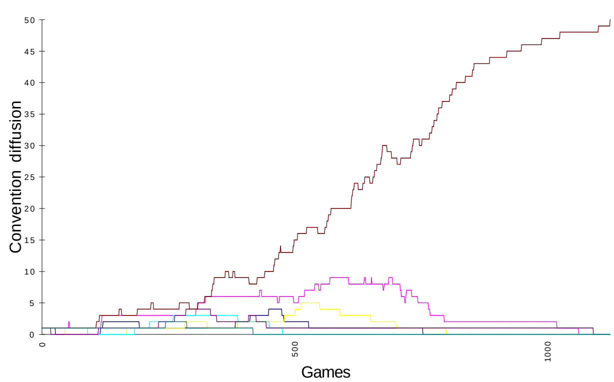

Figure 1: Word competition in a population of 50 agents. Each curve represents the diffusion of a word convention in the population. Eventually, one convention dominates being used by the 50 agents (Exp 1.1)

Implicit evaluation

Model 1.2. We now consider a set of words of unequal qualities. For instance, some are more resistant to noise. We model that in a very crude way by associating to each convention a mutation probability $P_m(W_i)$ between 0 % and 100%. For each game, a random test is done to check whether the word has been transmitted successfully or not. In case of failure, the word is transformed to another word randomly picked among all the possible ones.

For this experiment we choose a mutation probability that grows linearly with the word number. Thus for word W_i, the formula is:

$$P_m(C_i) = \frac{i}{\text{Number of words}} \tag{1}$$

Exp 1.2.a (N=50, 1000 runs) Figure 2 shows, for 1000 simulations, the distribution of the winning words for a population of 50 agents. Words with low mutation rates have been selected. An external observer could say that the agents are doing a *collective optimisation*. They are naturally converging towards the best words. The phenomena is based of a *implicit evalution* of the solution similar to the one described for forraging behavior in ant colony (Dorigo *et al.*, 1997). It means that the agent are not evaluating individually the quality of each word for choosing the more robust ones. Illadapted words simply mutate more often and cannot propagate as easily as the others.

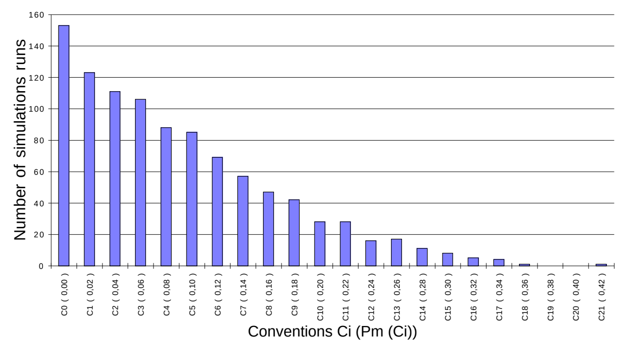

Figure 2: Distribution for 1000 simulation runs of the winning words for a population of 50 agents (Exp.1.2.a)

Reorganisation in the presence of an agent flux

In some cases though, the population might not converge towards the most robust words. The most important danger is *premature convergence*. If, for instance, a very good word appears in the population more lately during the experiment, it is probable that it will not be picked up because the positive feedback loop would have already caused the agents to converge towards a suboptimal one.

To consider an open population where agents are entering and leaving the population can correct this effect. Indeed, new agents entering the population have no special preference for the dominant word. They can discover the best solution and maybe, if it is really more robust than the one currently dominating, the outsider might eventually win.

Exp 1.2.b (N=50, different runs for different P_r). The following experiment is the same as the previous one, excepted that an agent flux, defined by the probability P_r of replacing an old agent by a new one, is applied. We want to see if this flux leads to a better selection of the words. For several values of P_r, we measure the proportion of simulation runs that end up with one of the three best word dominating. Beyond a certain value of P_r, the flux is too high to achieve convergence. The reorganisation can only be active near the edge of this threshold. Figure 3 shows this effect. We can draw an analogy between this effect and the role of temperature in optimisation techniques such as simulated annealing.

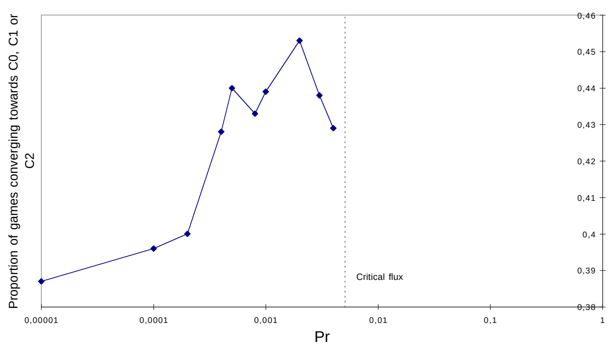

Figure 3: Proportion of simulation runs converging toward one of the three best solutions for different agent flux P_r (Exp 1.2.b)

In more complex models, an agent flux can also have a regularising effect. Because it increases the chance of picking up "good" linguistic conventions, conventions easier to learn will tend to be selected. Simon Kirby has illustrated this feedback loop on regularity in his simulation on the emergence of compositionality (Kirby, 1999b).

Conclusions

In this section we have identified the basic dynamics in the competition between different word conventions for naming one object.

- **Positive feedback loop**. If each agent is trying to induce the diffusion of each convention in the population, in order to use the one the most widely spread from its own point of view, then a positive feedback loop is created, leading to the domination of one convention. This dynamic is "blind" and does not prefer any convention *per se*.

- **Implicit evaluation of solutions**. But if some conventions are less easy to transmit, they will implicitely be left aside. Thus, best conventions tend to be chosen by the population.

- **Reorganisation and regularisation with an agent flux**. The presence of a flux of agent in the system avoids premature convergence. Better conventions (more robust, easier to learn) tend to be selected. The arrival of new agents enables a continuous parallel search for solutions that can replace the ones currently dominating. If needed it can cause a reorganisation in the communication system.

Competition during lexicon formation

In this section, we consider the case of the emergence of a lexicon: a mapping between a set of words and a set of objects.

Model 2.1. In this new model, the agents have to agree on names for a set of M objects. Each agent has an associated memory where are stored associations between words and objects. They use this memory to code an object into a word and to decode a word into an object. When several solutions are possible the agents choose the association with the highest score. Their associatiove memory is initially empty. Associations are progressively created as the agent interacts with other agents. As in the model of the previous section, a positive feedback loop enables lexicon self-organisation. Several experiments have shown that with such an architecture, a coherent lexicon emerges. Each word becomes associated to a single object and each object to a single name (Arita and Koyama, 1998; Steels, 1996; Steels and Kaplan, 1998b; Ferrer Cancho and Sole, 1998; Hutchins and Hazlehurst, 1995; Oliphant, 1997).

We consider a noisy environment where word transmission is difficult. Each word is an integer value between 0 and 1000. Each time a word is transmitted, a random number betwen $-B/2$ and $+B/2$ is added to the word. B is a measure of the global noise level. Each agent is equipped with a *filter* enabling him to select all the words in his associative memory which are at a distance D less than $D = B$. The structure of an interaction is the following:

1. The speaker randomly chooses an object o_1 between the different objects available and uses a word w_1 to name this object. If he doesn't have words associated with this object, the agent creates a new one (a random integer between 0 and 1000).

2. The word w_1 is transmitted to the hearer with an alternation between $-B/2$ and $+B/2$. The word heard is w'_1.

3. The hearer selects all the possible associations with a word close to w'_1 (at distance less than B). If no association is available, the speaker indicates what was the subject and the hearer creates a new association between w'_1 and the object o_1. If several associations are possible, the hearer chooses the one with the highest score: (w_2, o_2) .

4. if $o_1 = o_2$ the game is a success.

In case of success, the hearer increases the score of the association (w_2, o_2) with $+\delta$ and diminishes the score of competing associations (synonyms and homonyms) with $-\delta$. In case of failure, the hearer decreases the score of (w_2, o_2) with $-\delta$, the speaker indicates what was the subject and the hearer increases the score of the association (w_2, o_2) with $+\delta$, otherwise it creates it. Associations are initialy created with a 0 score. In the following experiment we take $\delta = 1$.

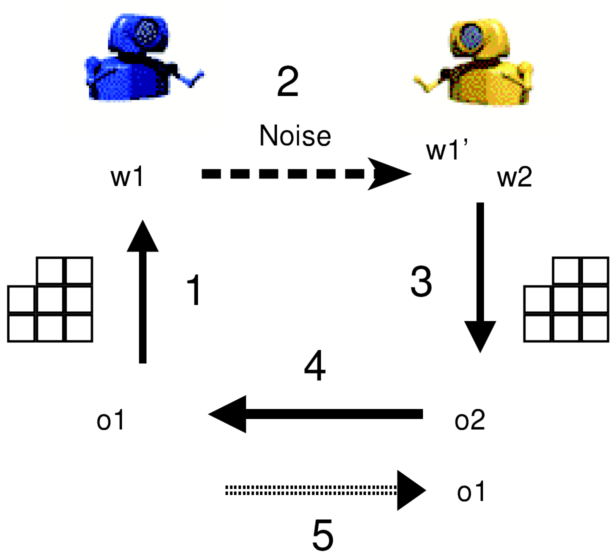

Figure 4: Interaction between agents using an associative memory (Model 2.1.).

Distinctivity

In the simple model of the previous section we have shown that collective dynamics lead to choose the "best" words to name an object. What are the best words in the current model ? A good word is a word than an agent will not confuse with another one that has a different meaning. A "good" lexicon should have set of words clearly distinct from one another depending on the object they name.

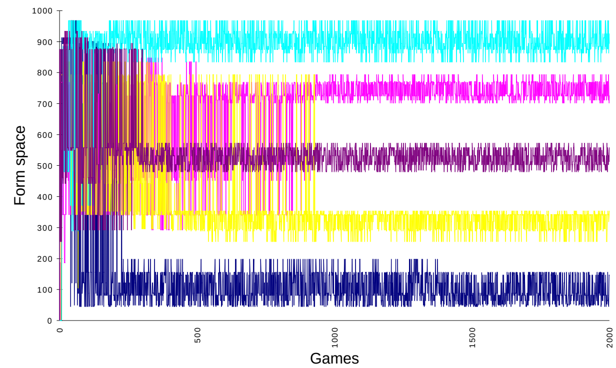

Figure 5: Evolution of the forms in the form space. After a first period of ambiguity, five well separated bands are forming to name each object (Exp. 2.1.a)

Exp 2.1.a $(N = 10, M = 5, B = 100)$ Figure 5 shows the evolution in the word space of the word associated with 5 objects in an experiment involving 10 agents. After an initial ambiguity period, five well separated bands in the word space are clearly indentifiable. Agents do not converge on a unique word form for each object. Each agent uses a different word. But as lexicon self-organisation is going on, these words tend to be very similar. For each object, they form a band in the word space which is cleary distinct form bands associated with other objects. No confusion is possible.

Figure 6 plots the same data as figure 5 showing the "average" word of each band. On this graph, it is easier to see the collective optimisation of distinctivity leading to a solution compatible with the level of noise present in the environment.

These results are somehow similar to the ones obtained by Bart de Boer (de Boer, 1997; De Boer, 1999). De Boer shows how the collective dynamics and noise lead a population of agents to converge towards a set of vowels optimally distributed in the phonological space in order to favor distinctiveness between them. Such emerging phonetic systems have high similarity with real ones as observed in natural languages.

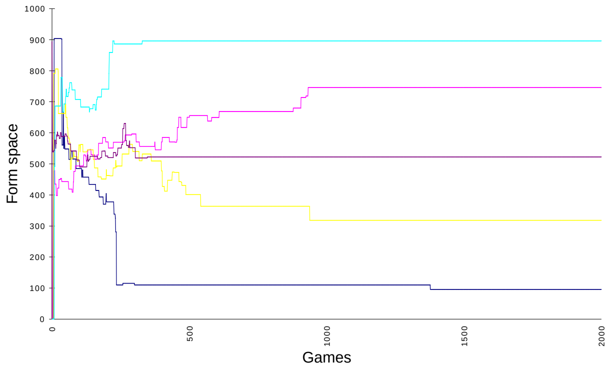

Figure 6: Evolution of the "average" forms in the form space (Exp. 2.1.a)

Compromise between distinctivity and robustness

Model 2.2 In the previously described experiments our model of a word - an integer - was very crude. In this section, each word is now a numeric chain of variable length. Each character of the chain is a number between 1 and 9. Noise is modelised by a probability of alteration P_m equal for each character. When a character mutates, it is simply replaced by a random character between 1 and 9.

As in the previous model, the hearer can look up in its lexicon for the chains that are "close" to the transmitted words. We define a distance D_c between word chains, similar to the traditional Hamming distance.

Let w_1 and w_2 be two words, the length of w_1 being either smaller or equal to the length of w_2. Let $w_1(i)$ and $w_2(i)$ be the character in position i in each of the chain. We define D_c as being the sum of the distance between the character of both chains to which is added 10 times their length difference, $l_2 - l_1$:

$$D_c(w_1, w_2) = \sum_i \|w_1(i) - w_2(i)\| + 10.(l_2 - l_1) \quad (2)$$

For instance the chains 1-4-5-2 and 1-4-5-7-3 are at a distance $5 + 10 = 15$. In the interaction the hearer selects the chains which are at a distance less than the threshold D.

We see that, with such a mechanism, too long or too short chains are naturally less adapted. Indeed, the longer a chain is, the more risk it has to be altered during transmission. In the first model, we have seen that such words generally loose the competition. But on the other side, if the lexicon is only composed of very short words, a single mutation might very often lead to confusion. A *compromise between word robustness and distinctivity* must be found: Short words are robust but easy to confuse, long words are easy to distinguish but difficult to transmit correctly.

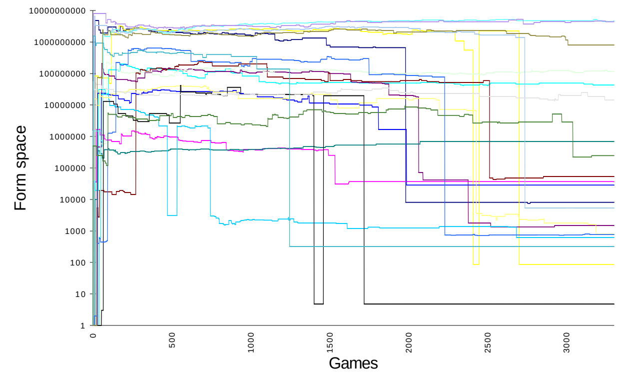

Figure 7: Example of evolution of the "average" words for 10 agents naming 20 objects with $D = 20$ and $P_m = 0.1$

Exp 2.2.a ($N = 10$, $M = 20$, $D = 20$, $P_m = 0.1$, 1 run) Figure 7 shows the evolution of average words for 10 agents naming 20 objects. As no word contains the "0" character, we can still visualize the situation in a one dimensional word space. In this representation, the values between 1 and 9 represent 1 word character, between 11 and 99, two word characters, etc. Considering this formalism, a logarithmic scale is appropriate. Each new division shows a new class of words. On the graph we observe that for this single run, as expected, the words of intermediary length constitute the majority of the final lexicon.

Exp 2.2.b ($N = 10$, $M = 20$, $D = 20$, $P_m = 0.1$, 100 runs) We have repeated experiment 2.2.a. a hundred times and analysed the distribution of all the words used by the agents after 5000 games (at this point, we have observed experimentally that the lexicon reaches a stable state). The results of the distribution of the word length are shown on figure 8. The distribution has a peak around words of length 3. Words too long or too short are less present in the final vocabularies.

Exp 2.2.c ($N = 10$, $M = 20$, $D = 5$, $P_m = 0.1$, 100 runs) The result of another series of experiments with a reduced noise tolerance level ($D = 5$) are shown on figure 9. The peak is now for words of length 2. As the noise level tolerance is reduced, a larger set of shorter words can be used as long as the tolerance is sufficient to cope with the noise level. It is the case in the conditions of these experiments.

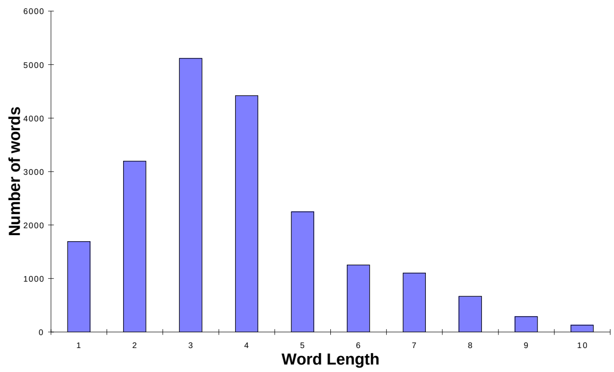

Figure 8: Distribution of word length on 100 simulation runs for 10 agents naming 20 objects with $D = 20$

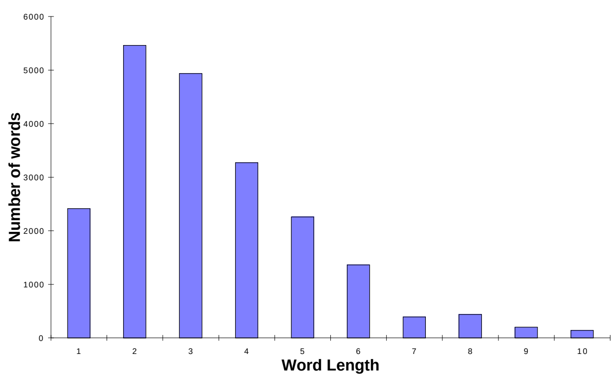

Figure 9: Distribution of word length on 100 simulation runs for 10 agents naming 20 objects with $D = 5$

Natural lexicon drift

We have seen with the model 2.1. that in a noisy environment, agents can converge on a stable system in which distinct bands in the word space are associated with distinct meanings. As we see in figures 5 and 6 this repartition in separated band does not evolve anymore once a stable solution has been found.

Exp 2.1.b ($N = 20, M = 2, B = 400, P_r = 0.01$) Graph 10 shows the evolution of the average form in the presence of an agent flux defined by a probability of replacing an old agent by a new one $P_r = 0.01$, for a population of 20 agents naming 2 objects. We see on the graph that the center of the bands are spontaneously evolving as new agents are entering the system. We will call this effect: the *natural lexicon drift*.

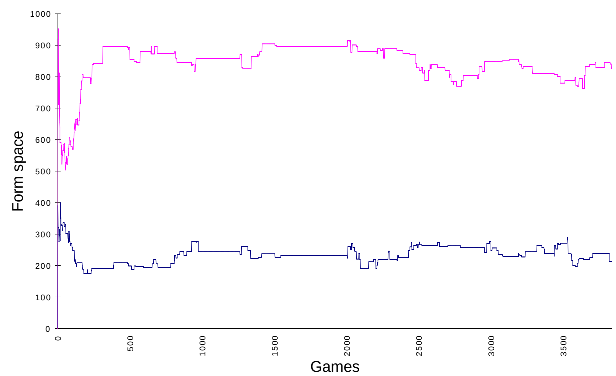

Figure 10: The natural lexicon drift. Spontaneous evolution of the "average" forms in presence of an agent flux (Exp 2.1.b)

This effect is easily understandable. A new agent tends to converge on words belonging to the existing bands for each meaning to express. But within this band, it has no reason to converge towards the exact center of the band. Thus the center is moving as the flux of new agents enters the system. The higher the agent tolerance on noise, the higher the amplitude of this drift.

These form bands are evolving spontaneously without any functional reasons. But this does not exclude that external pressures can direct these dynamics in a direction or another. The natural lexicon drift provides novelty and thus can lead to a more efficient reorganisation if needed.

We have shown in (Steels and Kaplan, 1998a) that this effect was active for more complex agent architectures including, like the simple agents of our model, a tolerance mechanims to cope with noise in the environment. Scott C. Stoness and Christopher Dircks have reproduced these results using another architecture based on neural networks (Dircks and Stoness, 1999). This shows that for such systems, collective dynamics are much more important than actual architecture and implementation details.

Conclusions

In this section we have explored the basic dynamics identified in the previous section with more complex models. Our conclusions are the following:

- **Distinctivity**. Noise during word transmission favors sets of words that are cleary distinct from one another when they mean different things. We experimentally observed the emergence of well separated bands in the word space. Each band is associated with a different word meaning.
- **Compromise between distinctivity and robustness**. When words can have different lengths, a compromise must be found between distinctivity and resilience to noise. Short words are easy to transmit but easy to confuse, long words are difficult to transmit correctly but are easily distinguishable from one another. We experimentally observe the convergence towards words of intermediary sizes.
- **Natural lexicon drift**. In the presence of noise and agent flux, we experimentally observe a spontaneous not functional lexicon evolution. This continuous exploration of the form space can lead to a more efficient reorganisation of the lexicon if needed.

Synthesis: the semiotic schemata

Other kind of competitions

We have considered in the previous sections very simple models in which meanings are simply discrete symbols without any particular properties. In some more complex architectures (Steels and Kaplan, 1999c; Steels and Kaplan, 1999b; Steels and Kaplan, 1999a) we have shown that when meanings are categories discriminating properties of the objects of the world, additional competitions can be observed. Some categories might be general and other specific. For instance, one might be used for describing a very particular shade of green, and another one for describing green objects in general. Depending on the environment and on the objects that need to be discriminated, sometimes general categories will be sufficient, sometimes specific ones will be needed.

Categories, like words, are competing with one another. Considering ompetition between isolated categories is not sufficient. The quality of a category needs to be evaluated regarding the category set to which it belongs. A specific category might survive if other categories are present to "back it up".

Associations linking words and meanings of different qualities can also be seen as competing units. A widely spread association has a real advantage, even on an association linking a very solid and easily distinguishable word with a very often used category.

In Simon Kirby's work, more complex system are competing with one another (Kirby, 1999b). He discusses the victory of a compositional system on a idiosyncratic one because the first one, being more regular, is easier to learn.

Eventually, several competitions can be observed at the same time, each of them involving part of languages of different sizes and types.

Semiotic schemata

Our point in this paper is to suggest that even though the kind of replicators involved in language emergence can be very different, the dynamics are merely always the same ones. These dynamics are the same as the one we have identified with the simple models of word competition that we have studied in the previous sections.

In (Steels and Kaplan, 1999b; Steels and Kaplan, 1999a), we have introduced the notion of *semiotic landscape* to analyse the complex dynamics involved in the emergence of word meanings. A semiotic landscape is a complex network linking objects, categories and words with associations of "different weight". If, for instance, the weight between a word and a meaning is strong in a semantic landscape its means that this associations is frequently observed in the agents behavior. All the cultural replicators, that we have identified in our experiment, can be seen as partial specification of a semiotic landscape. Words and meanings are simple nodes. Associations are couple of nodes and their links. Lexicons are more complex configurations.

Analysing what should be the right selection unit for studying artificial genetic dynamics, John Holland has introduced the notion of *schema* (Holland, 1995). A schema is a partial specification of the genome. By analogy, we introduce the notion of *semiotic schema* as a partial specification of a semiotic landscape. A word is a semiotic schema, a meaning is a semiotic schema, associations, sets of distinct words and even set of associations are semiotic schemata.

Order of a semiotic schema

A genetic schema, in artificial systems like genetic algorithms, can be modeled as a chain using only three characters 0, 1 and * (Holland, 1995). Each character of the chain can be seen as a particular gene which can take two values 0 or 1 or be undetermined, in which case it takes the value *. For instance, if the length of the genome is limited to $L = 10$ characters, $s = (*, *, *, 1, *, *, 0, *, *, *)$ is a possible schema, specifying only the genes number 4 and 7.

Schemata can be compared by studying their spreading in the population. For instance, if the schema s is present in all the population members, it means that all the genomes have the value 1 for the fourth gene and the value 0 for the seventh one. *Landscapes* are representations of the distribution of schemata at a given state of the evolution. For each possible schema, the number of agents in which this schema is present can be plotted. In practice, such landscapes are difficult to draw as the number of possible schemata can be very high. For a genome of length L, 2^L different schemata are possible. Yet, the landscape metaphor is a good starting point to visualize the competition between schemata.

Unfortunately, the situation for semiotic schemata is a bit more complex. It seems that several kinds of competitions are going on in parallel involving schemata of different complexity. We can introduice the notion of *order* of a schema. Words, groups of words, meanings and groups of meanings are first order semiotic schemata. Associations and lexicons are second order semiotic schemata. For a given order, semiotic schema can be modeled exactly like genetic schema using a chain composed with the characters 0, 1 and *.

For instance, the results of Exp 1.2.a, where 50 words of decreasing quality were in competition, can be seen as the competition between schema of Length $L = 50$ where each character codes for a word in the word space. Figure 2 shows that schemas starting with "1" in the first positions have a higher fitness than others. Results of Exp 2.1.a can be seen as the competition between schema of length $L = 1000$ and show that schemata including equidistributed "1" in the word space have a higher fitness than schemata where "1" are close from one another. Results of Exp 2.2.b show that schemata including words of intermediary length have a higher fitness than others.

The same analysis can be done for schemata of order 2. In these schemata, each character is a possible association of the lexicon. The length L of these schemata is equal to the product of the number W of possible words and the number M of objects to name. Once a stable mapping has been found, the semiotic landscape of such systems is defined by M distinct peaks corresponding the the M objects to name.

Selection dynamics

Semiotic schemata are replicators. The more complex they are, the more difficult it is for them to replicate. Their competition progressively structures the semiotic landscapes defining the common lexicon which is emerging. In particular kinds of word competition that we have studied in this paper, we observed three kinds of selection dynamics:

- **Individual choices select good schemata.** An agent has a way of evaluating semiotic schemata. The agent will use the ones that have proved to be efficient for communicating in past interactions. In the models we presented, a score was monitoring the success and the failure of each association and thus indirectly measuring their diffusion in the population. This dynamics create a positive feedback loop leading some semiotic schemata to be used more and more often. For this individual selection, schemata that are widely spread, resilient to noise and easy to learn have a selective advantage.

- **Agent flux ensures regularisation and reorganisation.** Individual selection is responsible of the lock-in effect on particular schemata. Because of premature convergence, the schemata chosen might not be the most efficient to communicate. The presence of a constant agent flux in the system puts additional pressure on the system for selecting really good schemata. An efficient schema might have been constructed by some individuals but appeared, later on, to be too difficult to transmit culturally to each new generation of agents. This agent flux creates a positive feedback loop on simplicity and therefore on regularity. The more regular and easy to learn a schema is, the more likely it is to pass the "generation bridge". The agent flux is also responsible for a continuous exploration of new possible schemata. Newborn agents might find simpler and more efficient solutions. If they are really good schemata they might replace the existing dominant ones.

- **Neutral dynamics ensures spontaneous novelty**. We have also observed some neutral dynamics. Schemata might be victims of neutral drifts similar by some aspects to process described in neutralist evolutionary theories (Kimura, 1983). Neutral dynamics are observed when a small level of noise causes inter-individual variations for schemata and new agents are regularly entering the population.

We guess that these dynamics are the most important for a large kind of semiotic schemata. But this remains to be tested in future works.

The role of noise

Semiotic schemata are always used in a particular context of a given environment. In the model we presented the effect of environments was limited to the addition of noise during the transmission phase. Noise has a double effect on semiotic schemata:

- **Noise as a diversity generator.** In artificial genetic evolution, noise could be assimilated to different mutations and errors that can appear during the copying phase of the genetic schemata. Noise is a diversity generator. In our dynamics also, noise could be a source of novelty for the creation of new semiotic schemata.

- **Noise as a pressure for selecting good schemata**. But its most important role is in the destabilisation of ill-adapted schemata. Words not distinct enougth from one another could not survive in the presence of noise. Too long words are avoided. Only robust categories that are efficient in noisy environments are selected, etc.

Adaptation not optimisation

At the beginning of the experiments, pool of semiotic schemata are unstructured. Then, the dynamics select sets of schemata that are well adapted to the environment in which the agents are communicating. During this process, we might be tempted to say that the "quality" of the schemata increases. But, like for species natural evolution, optimisation stops once adaptation is reached. We have seen that in the presence of noise, well separated bands of words were emerging. Though, once a stable solution was found, this optimisation of distinctivity stops. This effect has been also observed in more complex architectures where residual polysemy was observed (Kaplan, 2000). In all these situations, there is no absolute optimisation, only the search for stable solutions adapted to the environment. Once a set of stable schemata emerges, it can be considered as a higher level schema that might enter in competition with other higher level schemata.

Conclusion

The understanding of cultural dynamics involved in the emergence of communication systems is only at its beginning. In this paper, we have shown on simple models some basic mechanisms that organise the selection of words in the emergence of a lexicon. We believe that these mechanisms apply to a larger set of replicators that we call semiotic schemata. Adapted semiotic schemata are culturally selected through individual choices in a population continuously renewed. As this process goes on, shared communication systems emerge.

Acknowledgment

This research was carried out at in the Language Group of the Sony Computer Science Laboratory in Paris. I thank Luc Steels, Angus McIntyre, Jelle Zuidema and Sébastien Picault for precious comments on this work.

References

T. Arita and Y. Koyama. Evolution of linguistic diverstity in a simple communication system. In C. Adami, R. Belew, H. Kitano, and C. Taylor, editors, *Proceedings of Artificial Life VI*, pages 9–17. MIT Press, June 1998.

J. Batali. Computational simulations of the emergence of grammar. In J. Hurford, C. Knight, and M. Studdert-Kennedy, editors, *Approaches to the Evolution of Language: Social and Cognitive bases*, pages 405–426. Cambridge University Press, Cambridge, 1998.

S. Blackmore. *The meme machine*. Oxford Univesity Press, 1999.

A. Darmesteter. *La vie des mots*. Delagrave, Paris, 1937.

R. Dawkins. *The selfish gene*. Oxford University Press, 1976.

R. Dawkins. Replicator selection and the extended phenotype. In E. Sober, editor, *Conceptual issues in evolutionary biology*. The MIT Press, Cambridge, Ma., 1984.

B. de Boer. Generating vowel systems in a population of agents. In P. Husbands and I. Harvey, editors, *Proceedings of the Fourth European Conference on Artificial Life*, Cambridge, MA, 1997. MIT Press.

Bart De Boer. *Self-organizing phonological systems*. PhD thesis, VUB University, Brussels, 1999.

Daniel Dennett. *Darwin's Dangerous Idea. Evolution and the Meaning of Life*. Simon and Schuster, New York, 1995.

C. Dircks and S. Stoness. Effective lexicon change in the absence of population flux. In D. Floreano, J-D Nicoud, and F. Mondada, editors, *Advances in Artificial Life (ECAL 99)*, Subserie of Lecture Notes in Computer Science, pages 720–724, Cambridge, MA, 1999. Springer-verglag.

M Dorigo, V. Maniezzo, and A. Colorni. The ant system: optimisation by a colony of cooperating agents. *Transactions on Systems, Man and Cybernetics Part B*, 26(1):29–41, 1997.

R. Ferrer Cancho and R. Sole. Naming games through distributed reinforcement. manuscript, 1998.

J. Holland. *Hidden order: How adaptation builds complexity*. Addison-Wesley Publishing Company, 1995.

J. Hurford. Social transmission favours linguistic generalisation. In C. Knight, J. Hurford, and M. Studdert-Kennedy, editors, *The emergence of Language*. To appear, 1998.

E. Hutchins and B. Hazlehurst. How to invent a lexicon: the development of shared symbols in interaction. In N. Gilbert and R. Conte, editors, *Artificial Societies: The Computer Simulation of Social Life*. UCL Press, 1995.

F. Kaplan. *L'émergence d'un lexique dans une population d'agents autonomes*. PhD thesis, LIP6 - Université Paris VI, 2000.

M. Kimura. *The neutral theory of molecular evolution*. Cambridge University Press, 1983.

S. Kirby. *Function, selection and innatensess: The emergence of language universals*. Oxford University Press, 1999.

S. Kirby. Syntax out of learning: the cultural evolution of structured communication in a population of induction algorithms. In D. Floreano, J-D Nicoud, and F. Mondada, editors, *Advances in Artificial Life (ECAL 99)*, Subserie of Lecture Notes in Computer Science, pages 694–703, Cambridge, MA, 1999. Springer-verglag.

M. Oliphant. *Formal approaches to innate and learned communicaton: laying the foundation for language*. PhD thesis, University of California, San Diego, 1997.

L. Steels and F. Kaplan. Spontaneous lexicon change. In *Proceedings of COLING-ACL 1998*, pages 1243–1249, Montreal, August 1998. ACL.

L. Steels and F. Kaplan. 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 1998. MIT Press.

L. Steels and F. Kaplan. Bootstrapping grounded word semantics. In T. Briscoe, editor, *Linguistic evolution through language acquisition: formal and computational models*. Cambridge University Press, 1999.

L. Steels and F. Kaplan. Collective learning and semiotic dynamics. In D. Floreano, J-D Nicoud, and F. Mondada, editors, *Advances in Artificial Life (ECAL 99)*, Subserie of Lecture Notes in Computer Science, pages 679–688, Cambridge, MA, 1999. Springer-verglag.

L. Steels and F. Kaplan. Situated grounded word semantics. In *Proceedings of IJCAI 99*, pages 862–867, Stockholm, 1999.

L. Steels. Self-organizing vocabularies. In Chris Langton, editor, *Proceeding of Alife V*, Nara, Japan, 1996.

L. Steels. The synthetic modeling of language origins. *Evolution of Communication Journal*, 1(1):1–34, 1997.