A new approach to class formation in multi-agent simulations of language evolution

Frédéric Kaplan
SONY Computer Science Laboratory - 6 Rue Amyot, Paris
LIP6 - OASIS/MIRIAD - UPMC - 4, place Jussieu F-75252 Paris
kaplan@csl.sony.fr

Abstract

Multi-agent models of language evolution usually involve agents giving names to internal independently constructed categories. We present an approach in which the creation of categories is part of the language formation process itself. When an agent does not have a word for a particular object it is allowed to use the existing name of another object, close to the original one as defined by an analogy function. In this way, the names in the shared lexicon that has evolved in a collective way, directly yield the different object classes. We present the results of several simulations using this model showing under what conditions the agents will develop meaningful classes. We also examine the effects of an influx and outflux of agents. Finally we discuss the prospects for models in which the classes would constitute relevant complex taxonomies.

1 Introduction

This work is part of a new approach on language origins and evolution. Instead of analyzing existing linguistic forms in order to trace back their roots, a bottom-up approach is considered. Language is viewed as a complex dynamical system that emerges through adaptive interactions between agents. Similar approaches are currently under study in social sciences [7][8]. The multi-agent paradigm seems well adapted to the study of these phenomena (see discussion in [6]). Interesting results have already been obtained for different areas of language: simple communication codes [2] [13] [15] [22], coordination [4] [5] meaning and class formation [16], conventional lexicons [10][17][20] , phonetics [3] and syntax [1][9][18]. An overview of this approach is given in [19].

This paper focuses on the links between class formation and lexicon building. This topic is a fundamental issue in cognitive science, linguistics and philosophy (see [12]). Most of the existing computational models study how agents can associate a single word with a referent [13][15][17][21][22]. The referents can be categories, classes of objects, concepts either predefined in the model or evolving from processes distinct from the language formation itself. This amounts to assuming that an object table corresponds to a concept 'table' independently of the agents naming it a "table". This is the idea of *perceptually grounded categories.*

It can be argued however that the categorization process is not, in fact, distinct from the language formation process. So it is of interest, in order to investigate the dynamic construction of meaning, to study models in which categories are a consequence of lexicon construction and evolution. We will call them *language-based categories.* The object of this paper is to present a simple model that implements this feature (Figure 1).

Our model is inspired by some of Wittgenstein's conceptions on lexicon formation and categorization [23]:

1. Linguistic structures emerge as individuals get involve in series of interactions called *language games.* For instance, the case of an individual A using the name of an object in order to have an individual B bringing it to him is a simple language game.

2. Categories are not necessarily defined by common properties. For instance, a category like *tool* would be difficult to define in terms of common features as there is a wide range of very different tools. Tools are more like members of a family, analogous to one another. Wittgenstein talks of *Family Resemblances.* To be called a tool, a thing

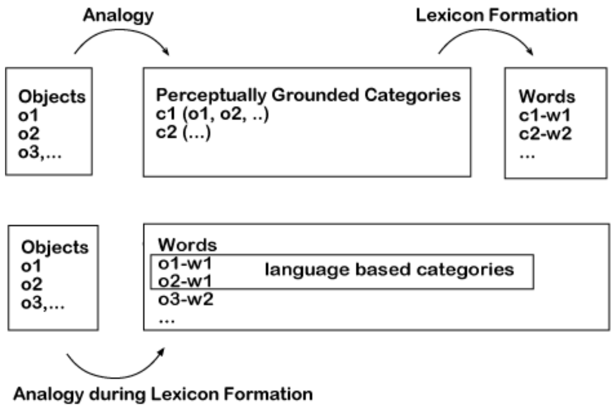

Figure 1. Two models of class formation

should be analogous to at least one member of the *tool* family.

3. Some members of a category are more representative than others; they are central. Wittgenstein gives the example of *numbers*, suggesting that the integers are central because, even though the concept of numbers might vary, integers are always considered as part of it. Many categories shows *membership gradience* and include central and peripheral members.

A possible approach to the formation of classes in general is to introduce, first, an *analogy distance* between objects and, second, a mechanism for clustering objects into classes according to their proximity in term of that analogy distance. In our work the mechanism for forming classes is the following: When agents do not have a word for a particular object they are allowed to scan their lexicon and use an existing name of a closely similar previously named object. This corresponds to babies using the word "dog" for any encountered animal.

When we run such a model, the collectively shared names directly yield the different objects classes. We are interested in exploring under different operating environment whether agents will develop a relevant set of classes.

Section 2 presents and illustrates the principles of these analogical naming games, a kind of language game derived from the naming games defined by Steels in [17]. Then a collection of measures and tools necessary to analyze the language and the classifications built by the agents is introduced with one example of lexicon formation. We present examples of the simulations using this new model, in particular investigating the effects of an influx and outflux of agents on classification building. We discuss briefly the prospects for models in which the classes would constitute relevant complex taxonomies and conclude by summarizing the interesting features of the model.

2 Self-organized vocabulary

2.1 Naming games

Previous work [17] has shown that a population of agents is able to adaptively build a set of shared conventions to name objects of their world. In these experiments, an agent is characterized by its lexicon. A lexicon consists of a set of associations between words and referents. It cannot be inspected directly by another agent. The *use* and the *success* of each association are monitored. Homonymy and synonymy may occur, since a single word can be associated with several objects and a given object can be associated with several words.
In the rest of the document we will consider the case in which a set of agents develops a shared vocabulary for identifying a given set of distributed objects that we will call referents. The agents interact with one another by playing *adaptive naming games*. During each game, two agents, a speaker and a hearer, are randomly selected from the population and interact as follows:

1. The speaker randomly selects an object as the topic of the interaction.

2. The speaker retrieves the word-referent associations where the referent is equal to the topic and picks the most preferred association according to a preference function.

3. The speaker communicates the word to the hearer.

4. The hearer retrieves the word-referent associations where the word is equal to the word used by the speaker. The game is successful if one of the decoded referents is the topic of the interaction.

When the game is not successful, each agent can adaptively modify its lexicon. If the speaker does not have a word to encode the topic, it may create a new word with a probability W_c and associate it with the topic. If the hearer does not know the word used by the speaker or does not associate it with the topic, it may acquire the new word-referent pair with a probability W_a.

The choice of the preference function is a crucial point. In the experiments described in this paper the agents always pick the word that has the best success/use ratio for a given referent, which means that they use the word-referent pair that has proven to be the most efficient for communication in the previous games.

A useful representation of the lexicon of an agent is a synthetic array in which each row represents a referent to be expressed (in our case, the different objects) and each column stands for a word in the agent's lexicon (similar representations are already used in [13], [15] and [22]). The values in the cells are the lexicon scores for each association (in our case, its normalized success/use ratio). Tables 1 and 2 show the lexicons of two agents, in a world where there are only 6 objects, after several games. To give illustrations of naming games, we can take the case of an interaction between speaker a1 and hearer a2. Several situations can happen depending on the topic chosen:

1. If the topic chosen is object1, a1 will say the word W1. a2 knows W1 is a possible name for this object (but it is not the word it would have used). The game is successful.
2. If the topic chosen is object4, a1 will say the word W2. a2 does not know this word. The game is not successful and a2 might add the new association in its lexicon.
3. If the topic chosen is object5, a1 does not have a word. The game is not successful and a1 might create a new word.

Table 1: Lexicon of agent a1

Referent	W1	W2	W3	W4	W5	W6	W7	W8	W9
object1	0.75			0.25					
object2	0.17				0.83				
object3			0.64			0.36			
object4		0.63					0.37		
object5									
object6									1.00

Table 2: Lexicon of agent a2

Referent	W1	W2	W3	W4	W5	W6	W7	W8	W9
object1	0.33			0.67					
object2					1.00				
object3			0.40			0.60			
object4							0.50	0.50	
object5						1.00			
object6									1.00

2.2 Analogical naming games

Analogical naming games are identical to standard naming games except that when an agent does not know the name of an object, it can scan its lexicon looking for the existing name of another object, close to the original one according to an *analogy function*. A word previously used for a single object can, by this mechanism, be *recruited* to name a group of objects.

In order to avoid making analogies when the number of words known by an agent is too small, a *Vocabulary Size Threshold* parameter (VST) fixes the minimum amount of words needed to begin the analogies. As we will show in section 4 the choice of this parameter is crucial.

In this paper, analogical naming games are studied within a very simple environment. The objects of the world are distributed in a 2D grid (such as the one of figure 2). The distribution is not completely random and contain some clustering. Of course, no information about this clustering is given to the agents. They have a global view of the objects spatial distribution. We choose to define an analogous object as the spatially closest one according to the Euclidiean distance. This means that in our experiments, the agents can name objects, not lexicalized yet, by the name of the spatially closest lexicalized object. As we show in the subsequent sections, the agents, by playing analogical naming games, build lexicons that can be seen as shared classifications where the objects are grouped into classes. In this model, relevant classes should grouped spatially closed objects. The Euclidiean distance is chosen for simplicity. Other distances can be used.

3 Measures and visualizations

In this section we present measures and visualizations to analyze the lexicons of the agents and therefore the classifications they build (a more complete study of possible measures of language evolution can be found in [11]) .

3.1 Communicative success and Coherence

Two basic measures are defined to study the global properties of the communication system built by the agents. The first one, *communicative success* $S_{\mathcal{A},k}$, is an indication of how much the agents succeed in communication, to what extent they understand each other. To plot its evolution, the ratio between the number of successful games and the number of games played in the population $\mathcal{A}$ is computed every k language games.

But the agents can understand each other without using the same lexicon. For instance, according to Table 1, A1 uses the word W1 to name object1. A2 understands this word as a possible name for object1 but would not use it if it was the speaker. In order to have an indication of how much the agents have converged towards a *single* vocabulary, it is necessary to define another measure: coherence.

The *coherence* $C_{\mathcal{A},t}$ of a lexicon in a group of agents $\mathcal{A}$ measures the average percentage of the most preferred words in the lexicon. For each referent the preferred word of every agent is gathered and the most commonly used preferred word is selected. Its percentage of use defines the coherence for this referent. The average coherence for all the referents is finally computed. Table 3 shows the computation of coherence in a population of 20 agents talking about 6 objects.

Table 3: Computation of coherence

Referent	Words	Coherence
object1	W1(15), W4(5)	15/20
object2	W5(17), W1(3)	17/20
object3	W3(10), W6(8), W2(2)	10/20
object4	W2(14), W7(4), W8(2)	14/20
object5	W6(19) W1(1)	19/20
object6	W9(20)	20/20
		0.79

We now consider a population of 20 agents playing analogical naming games in order to name 30 objects. The objects are spatially distributed in a 2D grid and clustered into 3 groups (C1 9 objects - C2 8 objects - C3 13 objects) as shown on figure 2.

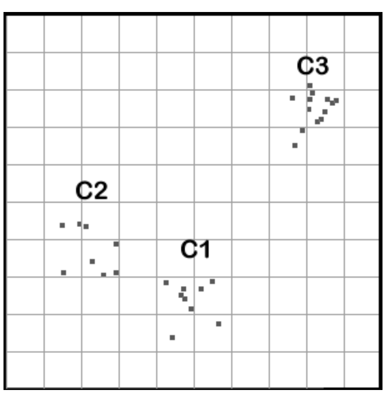

Figure 2. Spatial distribution of the objects

To prevent the agents from building too complex a vocabulary we limit the number of possible words to 5. Figure 3 shows the evolution of communicative success and coherence for the first 10000 games. After 5000 games the two curves have stabilized. The communicative success reaches its maximum value, which means that the agents fully understand each other. The high value of Coherence shows that the population has managed to build a shared vocabulary.

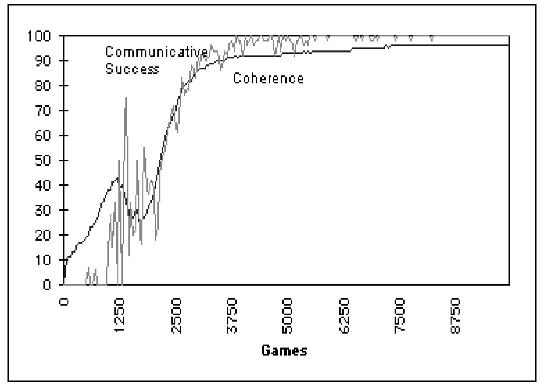

Figure 3. Evolution of communicative success and coherence

3.2 Global Lexicon

To analyze the shared vocabulary built by the agents, we need a synthetic representation of the set of individual agent lexicons. The same kind of arrays used for an agent in the previous section can be used again. For each referent the average lexicon score over the population is computed. Table 4 shows the global lexicon towards which the agents have converged after 10000 games in the experiment described in 3.1. We see, for instance, that W1 characterizes the agent of C1, W2 and W5 the agents of C2 and W3 and W4 the agents of C3.

Table 4 shows that some member seems central inside their categories. For instance among the objects called W5, the object13 seems the more representative as all the agents agree to name it this way. But, as the membership gradience remains rather difficult to see on the table, a special graphical monitor has been designed for that purpose.

3.3 Classification monitor

The Classification monitor has been designed to check graphically the correspondence between the global classification formed during the lexicalization process and the spatial distribution of the objects. When the global lexicon is represented in an array such as the one of Table 4, each category (objects with the same name) is characterized by a list of objects and

Table 4: Global lexicon built after 10000 games

Cluster	Referent	W1	W2	W3	W4	W5
C1	object1	1.00				
C1	object2	1.00				
C1	object3	1.00				
C1	object4	1.00				
C1	object5	1.00				
C1	object6	1.00				
C1	object7	1.00				
C1	object8	1.00				
C1	object9	1.00				
C2	object10		0.46			0.54
C2	object11		0.17			0.83
C2	object12		0.96			0.04
C2	object13					1.00
C2	object14		0.27			0.73
C2	object15		0.17			0.83
C2	object16		0.98			0.02
C2	object17		1.00			
C3	object18				1.00	
C3	object19				1.00	
C3	object20				1.00	
C3	object21			0.22	0.78	
C3	object22				1.00	
C3	object23				1.00	
C3	object24				1.00	
C3	object25				1.00	
C3	object26			1.00		
C3	object27			1.00		
C3	object28			0.33	0.67	
C3	object29			1.00		
C3	object30			1.00		

the lexicon scores W_o associated with each of them. For each category three quantities can be computed:

1. The Mass:

$$M = \sum_{o=1}^{N} W_o \qquad (1)$$

2. The center of gravity:

$$G \begin{pmatrix} x_g = \frac{1}{M} \cdot \sum_{o=1}^{N} W_o.x_o \\ y_g = \frac{1}{M} \cdot \sum_{o=1}^{N} W_o.y_o \end{pmatrix}$$

3. The mean quadratic distance from the center of gravity:

$$D = \sqrt{\frac{1}{M} \sum_{o=1}^{N} W_o.distance(G, O)^2} \qquad (2)$$

N: Number of objects in the category, W_o: lexicon score of the object o, (x_o, y_o) : coordinates of the object o.

G can be viewed as the ideal prototype of a category and D shows how precise the word is. For instance, a large D means that the word is used for an important number of objects. For each word of the global lexicon a circle of center G and of radius D is drawn. The visualization of the lexicon described in Table 4 is shown on figure 4. It gives a more synthetic view of the classification and shows centrality and membership gradience for each category. In particular we see that W2 and W5 are not synonyms and define two subgroups of C2.

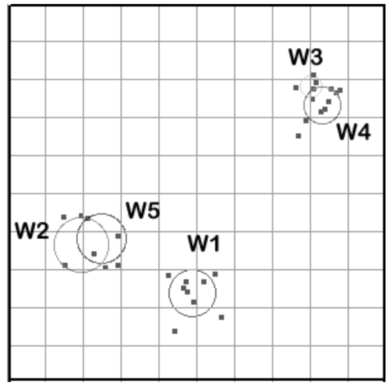

Figure 4. Graphical representation of the correspondence between the built classification and the spatial distribution of the objects

3.4 Classification quality

The Classification monitor enables us to graphically compare two classifications. But this comparison remains subjective. In order to have more valuable data, we define a measure called *Classification Quality*.

$$CQ = 1 - \frac{D_{mean}}{D_{total}} \qquad (3)$$

$$D_{mean} = \frac{1}{Nw} \sum_{w=1}^{Nw} \sqrt{\frac{1}{Mw} \sum_{o=1}^{No} Wow.distance(Gw, O)^2} \qquad (4)$$

$$D_{total} = \sqrt{\frac{1}{No} \sum_{o=1}^{No} distance(G, O)^2} \qquad (5)$$

No: Number of objects, Nw: Number of words, Wow: lexicon score for the association o-w, Mw: Mass associated with the word w, Gw: center of gravity associated with the word w, G: center of gravity of all the objects.

The mean radius of the circles gives an idea of the precision of the classification. If we consider two classifications using the same number of words, the one with the smaller mean radius would be the more relevant. For the classification of Table 4 and Figure 4, $CQ = 1 - 0,17 = 0,83$.

4 Experiments

In all the experiments described in this section, the objects are distributed into a 2D Grid. We consider a population of 20 agents. The word creation rate Wc is chosen 0.1 and word acquisition rate Wa is always 1 which means that the agents learn a new word-referent association the first time they encounter it. For the first two experiments, the number of possible words (and therefore of possible categories) is limited to 5 and accidental homonymy is prohibited, which means that two agents cannot create the same word. In the last one, we give an example of a simulation with an unlimited set of possible words.

4.1 The choice of the VST parameter

The Vocabulary Size Threshold parameter is the number of words needed before an agent starts to make associations based on analogy between objects. If an agent starts to make associations when it has only a few words, it might call all the objects more or less the same way which would lead to uninteresting classifications. Figure 5 compares in an example the classifications made after 5000 games for different values of the VST parameter. The classification is far more accurate for the higher value of the VST parameter.

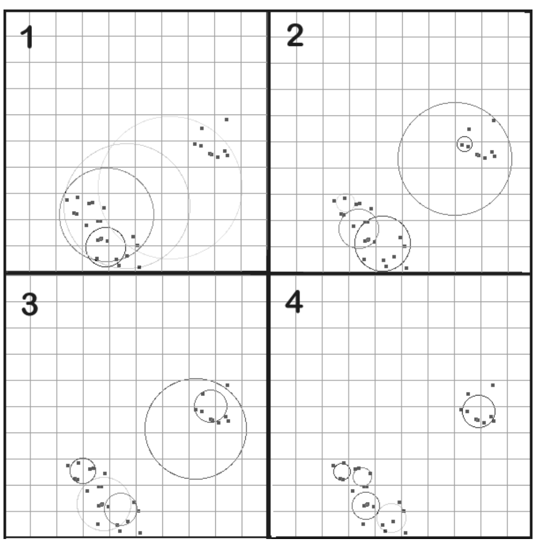

Figure 5. Classifications gets more relevant as VST increases

This result is confirmed by the computation of the Classification Quality measures over 20 simulations with different distributions of the objects.

VST parameter	Mean CQ
1	0.37
2	0.62
3	0.64
4	0.70

4.2 Classification refinements due to influx and outflux of agents

We now consider the case of an open system where there is an influx and outflux of agents. The population is kept at the same size as every time a new agent enters the system, another is randomly suppressed. When an agent enters the system, its lexicon is empty. Its entrance has two effects:

1. It is a source of novelty because it might create new word-referent pairs and discover associations previously ignored by the agents of the population.

2. It is a source of stability because as a new agent interacts equally with all the agents of the population, its lexicon is a "snapshot" of the global lexicon of the population.

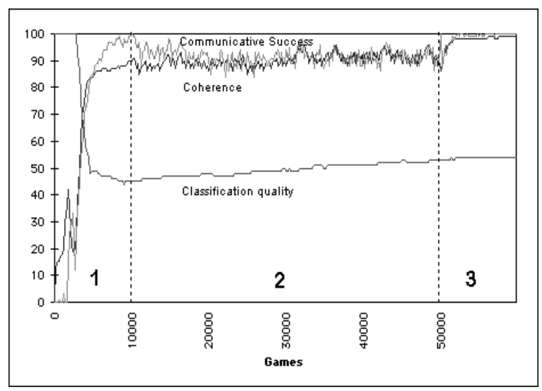

Figure 6. Effects of an influx and outflux of agents on communicative success, coherence and classification quality

Figure 6 shows an example of the evolution of communicative success, coherence and classification quality for a population of 20 agents talking about 50 objects organized in 5 large overlapping clusters. The simulation is divided in three parts

1. First the system is kept closed for 10000 games. Communicative success and coherence reach a high value. A shared vocabulary emerges.

2. After 10000 games the system is opened and an agent is changed every 500 games during the next

40000 games. Each time an agent is changed, communicative success and coherence drop for a while and then recover approximately its previous value. The regular renewal of the population can be seen as a series of *perturbations* with which the language has to cope.

3. After 50000 games the system is closed again. Communicative success and coherence rise to very high values again.

The effect of the influx and outflux of agents on the classification previously built is of great interest. Figure 7 is a comparison between the classification after 10000 games and the final clusters reached after 60000 games.

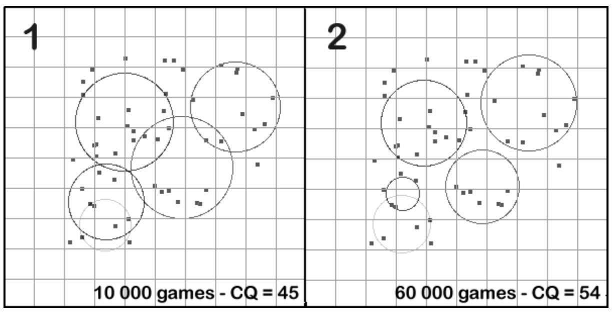

Figure 7. Classification refinements due to influx and outflux of agents

According to the classification quality measures the final classification is significantly more relevant. Figure 6 shows that CQ regularly increases as agents are entering and leaving the population. As new agents learn and reinvent the language, irrelevant associations tend to disappear and little by little the classification becomes more accurate.

4.3 Taxonomy formation

We have for the moment only studied experiments where the number of possible words was limited. This constrained environment made the simulation analysis easier. In this last experiment the number of words is unlimited as it is usually the case in our simulations of language evolution. Figure 8 shows the classification built by 20 agents for naming 50 objects. The VST parameter is set to 10. 77 words have been invented. As shown on Figure 9 one half of them are used to name individual objects, the other half for groups of objects. One object is often named with several words: its name, the name of the small group of its neighbors, the name of the cluster it belongs to, etc. From this point of view, the lexicon can be seen as a taxonomy that classifies the objects in a hierarchy of categories.

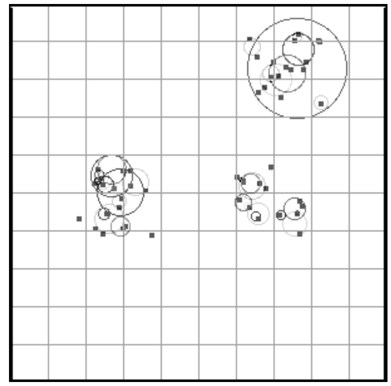

Figure 8. Classification made by 20 agents for 50 objects after 40000 games

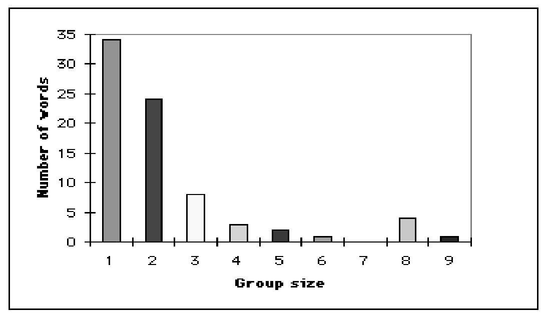

Figure 9. Number of words for each group size in the final vocabulary

5 Conclusion

This paper presents how a group of agents can form language-based categories through a basic interaction model for sharing lexicon conventions. This model, deliberately kept simple, presents very interesting features:

1. This approach does not assume any kind of preexisting dedicated structure to store categories. All the knowledge of the agents is contained in their lexicon, a set of associations between objects and word-forms. But as analyzing the different possible names for a given object is a way to characterize it, we believe that the agents can use this knowledge as efficiently as internal categories.

2. As categories are constructed in a collective way through language interactions, no central agency is needed. Agents entering the population learn the current naming conventions and in this way build a categorization system. As the population change, the global classification tends to be refined and to become more and more accurate.

3. Language-based categories share common characteristics with current views of categorization in cognitive science: Family resemblances, Centrality and Membership gradience. They are especially well adapted for representing fuzzy and evolving classifications.

We believe that this new approach, either used alternatively or together with the perceptually grounded one, can open interesting new perspectives in the building of multi-agent simulations of language evolution.

6 Acknowledgments

This work was done at Sony Computer Science Laboratory in Paris in collaboration with the Paris VI University. I thank Luc Steels and Alexis Drogoul for their valuable comments on my work. All the simulations of this paper have been constructed with the BABEL software developed by Angus McIntyre at Sony CSL-Paris [14]. This toolkit has proven to be very efficient for setting up new simulations rapidly.

References

[1] J. Batali. Computational simulations of the emergence of grammar. In J. Hurford, C. Knight, and M. Studdert-Kennedy, editors, *Evolution of Human Language*. Edinburgh University Press, Edinburgh, 1998.

[2] A. Cangelosi and D. Parisi. The emergence of language in an evolving population of neural networks. In *Proceedings of the 18th Conference of the Cognitive Science Society*, San Diego, 1996.

[3] 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*. MIT Press, 1997.

[4] E. de Jong. Multi-agent coordination by communication of evaluations. In M. Boman and W. Van de Velde, editors, *Proceedings of the 8th European Workshop on Modelling Autonomous Agents in a Multi-Agent World*, Berlin, 1997. Springer-Verlag.

[5] E. Di Paolo. An investigation into the evolution of communication. *Adaptive Behavior*, 2(6), 1998.

[6] A. Drogoul and J. Ferber. Multi-agent simulation as a tool for modeling societies: Application to social differentiation in ant colonies. In *Decentralized A.I. 4*. North-Holland, 1992.

[7] J. Epstein and R. Axtell. *Growing Artificial Societies: Social Science from the Bottom Up*. MIT Press, Cambridge, Ma, 1995.

[8] N. Gilbert and R. Conte. *Artificial Societies. The computer simulation of social life*. UCL Press, 1995.

[9] T. Hashimoto and T. Ikegami. Emergence of net-grammar in communicating agents. *BioSystems*, 38:1–14, 1996.

[10] 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.

[11] F. Kaplan, A. McIntyre, and L. Steels. Measures and visualizations of language evolution, 1998. Submitted to Publication.

[12] G. Lakoff. *Women, Fire and Dangerous Things: What Categories Reveal about the Mind*. Chicago Press, Chicago, 1987.

[13] B. MacLennan. Synthetic ethology: An approach to the study of communication. In C. Langton, editor, *Artificial Life II*, Redwood City, Ca., 1991. Addison-Wesley Pub. Co.

[14] A. McIntyre. Babel: A testbed for research in origins of language, 1998. Submitted for Publication.

[15] M. Oliphant and J. Batali. Learning and the emergence of coordinated communication. *Submitted to Cognitive Science*, 1996.

[16] L. Steels. Perceptually grounded meaning creation. In M. Tokoro, editor, *Proceedings of the International Conference on Multi-Agent Systems*, Cambridge, Ma, 1996. The MIT Press.

[17] L. Steels. Self-organizing vocabularies. In C. Langton, editor, *Proceeding of Alife V*, Nara, Japan, 1996.

[18] L. Steels. 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, 1997. Morgan Kauffman Publishers.

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

[20] 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.

[21] P. Vogt. Perceptual grounding in robots. In A. Birk and J. Demiris, editors, *6th European Workshop on Learning Robots*, 1997.

[22] G. M. Werner and M. G. Dyer. Evolution of communication in artificial organisms. In C. G. Langton, C. Taylor, and J. Farmer, editors, *Artificial Life II, Vol.X of SFI Studies in the Sciences of Complexity*, Redwood City, Ca., 1991. Addison-Wesley Pub.

[23] L. Wittgenstein. *Philosophical Investigations*. Macmillan, New York, 1953.