UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

## Title
How Nouns and Verbs Differentially Affect the Behavior of Artificial Organisms

**Permalink**
https://escholarship.org/uc/item/34n727d4

**Journal**
Proceedings of the Annual Meeting of the Cognitive Science Society, 23(23)

**ISSN**
1069-7977

**Authors**
Cangelosi, Angelo
Parisi, Domenico

**Publication Date**
2001

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

# How Nouns and Verbs Differentially Affect the Behavior of Artificial Organisms

**Angelo Cangelosi (acangelosi@plymouth.ac.uk)**
PION Plymouth Institute of Neuroscience and Centre for Neural and Adaptive Systems
University of Plymouth
Drake Circus, Plymouth, PL4 8AA, UK

**Domenico Parisi (parisi@ip.rm.cnr.it)**
Institute of Psychology, National Research Council
Viale Marx 15, Rome, 00137, Italy

**Abstract**

This paper presents an Artificial Life and Neural Network (ALNN) model for the evolution of syntax. The simulation methodology provides a unifying approach for the study of the evolution of language and its interaction with other behavioral and neural factors. The model uses an object manipulation task to simulate the evolution of language based on a simple verb-noun rule. The analyses of results focus on the interaction between language and other non-linguistic abilities, and on the neural control of linguistic abilities. The model shows that the beneficial effects of language on non-linguistic behavior are explained by the emergence of distinct internal representation patterns for the processing of verbs and nouns.

## Modeling the Evolution of Language

The recent development of computational evolutionary models (Wiles & Hallinan, in press) has contributed to the rebirth of interest in the origin and evolution of language. Computational models can directly simulate the evolution of communication and the emergence of language in populations of interacting organisms (Cangelosi & Parisi, in press; Dessalles & Ghadakpour, 2000; Steels, 1997). Various simulation approaches are used such as communication between rule-based agents (Kirby, 1999), recurrent neural networks (Batali, 1994; Ellefson & Christiansen, 2000), robotics (Kaplan, 2000; Steels & Vogt, 1997), and internet agents (Steels & Kaplan, 1999).

Artificial Life Neural Networks (ALNN) are neural networks controlling the behavior of organisms that live in an environment and are members of evolving populations of organisms. ALNN models have been used to simulate the evolution of language (Cangelosi & Parisi, 1998; Cangelosi, 1999; Cangelosi & Harnad, in press; Parisi, 1997). For example, in Cangelosi and Parisi's (1998) model organisms evolve a shared lexicon for naming different types of foods. Communication signals are processed by neural networks with genetically inherited connection weights and the signals evolve at the population level using a genetic algorithm with no changes during an individual's lifetime.

ALNN models provide a unifying methodological and theoretical framework for cognitive modeling because of the use of both evolutionary and connectionist techniques and the interaction of the organisms with a simulated ecology (Parisi, 1997). All behavioral abilities (e.g., sensorimotor skills, perception, categorization, language) are controlled by the same neural network. This unified framework permits the study of various factors affecting language evolution, such as the differences between genetic and learned communication systems, the adaptive role of both simple and compositional languages, the neural control of language, the reciprocal influences between language and cognition.

### Emergence of compositional languages: verbs and nouns

The evolutionary emergence of messages that combine two linguistic signals has been studied with ALNN models. In Cangelosi and Parisi's (1998) model, organisms communicate using simple signals that are genetically inherited. In an extension of the model, word combination and language learning were introduced to simulate the emergence of compositional languages (Cangelosi, 1999; in press). The organisms' neural networks had two linguistic winner-takes-all output clusters so that two "words" were simultaneously uttered to name foods (different types of mushrooms). Parents acted as linguistic teachers of their offspring. Children learned to name foods by imitating their parents' descriptions using an error backpropagation algorithm.

The simulation results showed that about 60% of the populations evolved optimal languages, i.e., languages in which each category of food was correctly identified and named. In the remaining populations, only one category out of six was misclassified. Evolved languages were classified into three types: (1) Single-word, where the units in only one cluster are enough to differentiate all mushrooms; (2) Word-combination, where symbols from both clusters are needed to discriminate mushrooms; (3) Verb-Noun, where the units in one cluster are systematically associated with high-order categories (e.g., "verbs" for approaching/avoiding) and the other cluster is used for differentiating sub-categories (e.g., "nouns" for mushrooms of different color).

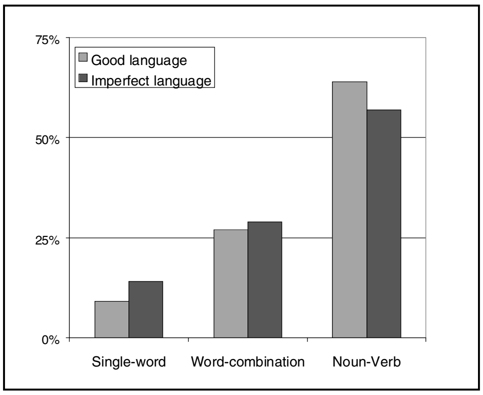


Figure 1: Distribution of languages in the 18 simulations with communication (at generation 400).

The distribution of language types (Figure 1) showed that there is a strong evolutionary tendency to evolve compositional languages, where the syntactic structure of messages reflects the hierarchical classification of mushrooms. In fact, the most frequent (e.g., 64% of good languages) combinatorial structure is that of predicate-argument, resembling a "verb-noun" sentence.

## Behavioral and Neural Factors in the Evolution and Acquisition of Language and Syntax

We will now focus on some issues about the acquisition and use of language, and on their relations with language evolution studies. These issues regard the interaction between language and other behavioral abilities, the stages of the acquisition and evolution of syntax, and the organization of neural representations in language processing. The first issue is quite an important and old one: How does language affect, and how is it affected by, other cognitive and behavioral abilities? Various language origin theories stress the importance of pre-existing sensorimotor knowledge for effective evolution of linguistic skills. For example, Rizzolatti and Arbib (1998) proposed a motor theory of language evolution based on imitation skills. Steels (2000) showed how his robotics models of language evolution support this theory. In Cangelosi and Parisi's (1998) ALNN model, they showed how language evolution relies on the evolution of basic cognitive abilities such as categorization. The dependence of language on previous sensorimotor skills, and the effects of language on this behavior will be looked at in the models presented here.

Researchers interested in both the evolution and the acquisition of language, are primarily concerned with the early stages of the development of linguistic abilities. In particular they focus on the transition from a non-linguistic stage where sensorimotor abilities dominate to a phase in which language and other high order cognitive skills emerge and take control of cognitive development. Although little empirical evidence is available for language evolution, data on language acquisition strongly support the conclusion that children learn nouns before verbs (Brooks & Tomasello, 1999). They handle nouns at around 18 months, while verbs are acquired later, from around 24 months. Verbs seem to follow a more gradual acquisition pattern, passing through an intermediate stage called "verb islands" (Tomasello, 1992). We will use data from our simulations to look for similar learning patterns in language evolution.

The investigation of the neural control of nouns vs verbs has been the focus of some interesting neuropsychological and brain imaging studies. For example, Caramazza and Hillis (1991) looked at the brain representation of noun and verbs in patients with brain lesions. Martin, Haxby, Lalonde, Wiggs & Ungerleider (1995) used PET to show that cortical sensory areas are active when the color word of an object is retrieved, while motor areas are involved in the processing of action words. ALNNs permit the investigation of internal representations involved in the processing of different syntactic classes such as nouns and verbs.

In the next section we will describe a new ALNN model of the evolution of syntax, specifically the verb-noun syntactic rule. This simulation will be use to study in detail the interaction between linguistic abilities and other behavioral and neural factors.

# Evolution of Verb-Noun Languages

The ALNN model described in Cangelosi, 1999 (cf. also Cangelosi, in press) showed a significant tendency to evolve compositional languages made up of verb-noun messages. To study the differences between verbs and nouns and how verb-noun languages affect and are affected by other behavioral, cognitive, and neural factors, a new model with a pre-defined compositional language will be used. The language includes four simple linguistic signals (words), two nouns and two verbs. Nouns are defined as linguistic signals that covary with the visual input. Verbs are defined as linguistic signals that covary with the action of the organism. Messages can include only a noun or only a verb or they can be a combination of a noun and a verb.

## The Model

The task used in the simulation is an object manipulation task (Schlesinger & Barto, 1999). At any given time the organism is grasping an object with its hand and it either pulls the object toward itself or it pushes the object away from itself. Two different objects are used, a vertical bar (object A) and a horizontal bar (object B). The object is perceived through a retina of 5x5=25 cells corresponding to 25 visual input units. The object occupies either three vertical cells or three horizontal cells in one of 9 possible locations in the retina. Hence, an object is encoded as a pattern of 25 bits with three 1s and twenty-two 0s. In addition to this visual input from the retina the organism's neural network receives a proprioceptive input encoding the current position of the organism's two-segment arm. This input is encoded in 4 input units, with units encoding proprioceptive information about the two pairs of muscles (extensor and contractor) of each of the two arm segments.

In the simulations with language the neural network includes 4 more input units encoding linguistic signals. Four linguistic signals are used, two nouns and two verbs, and they are localistically encoded in the 4 linguistic input units One noun designates the vertical object and a different noun designates the horizontal object. One verb designates the action of pushing and the other verb the action of pulling the object. In different occasions the organism can perceive only a noun or only a verb or both a noun and a verb.

There are two layers of hidden units that receive information from the input units and pass it to the 4 output units (Figure 2). The output units control the extension/contraction of the four arm muscles.

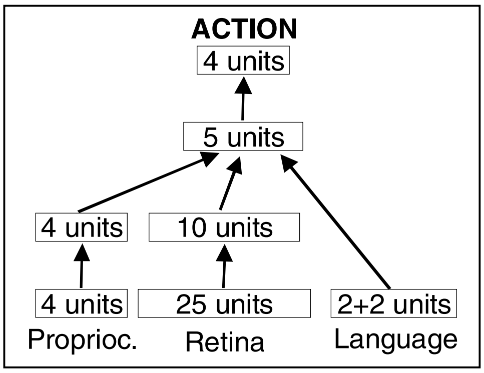


Figure 2 – The organism's neural network for the object manipulation task

The connection weights allowing the neural network to appropriately manipulate the two objects are developed using a genetic algorithm. At the beginning of a simulation 80 genotypes are randomly generated each encoding the connection weights of a single individual. These 80 individuals constitute the first generation. The 20 best individuals are selected for reproduction, with each individual generating 4 offspring with the same genotype (connection weights) of its single parent except for the addition of some random changes to some of the weights (random mutations). The process is repeated for 2000 generations.

Three experimental conditions were used. In the first condition, called "No-Language", an organism lives for a single epoch consisting of a total of 360 input/output mappings or moves (2 object types x 9 positions x 20 moves per task). Only the retina and the proprioceptive information are provided as input to the network. When the organism sees object A, it always has to push it away from itself; when it sees object B, it has to pull it towards itself. The fitness formula computes the total number of tasks successfully completed.

The second experimental condition is called "Late-Language". At generation 1000 a copy of the populations of the No-Language condition is made. From this generation onwards the organisms have a longer lifetime and they are exposed to language. Ten new epochs with language are added to an individual's lifetime, which therefore now includes 11 epochs, 10 with language and 1 without language. In 5 of the linguistic epochs an individual receives both the linguistic input and the retina and proprioceptive inputs, whereas in the remaining 5 epochs only the linguistic input and the proprioceptive input are present and the retina input is shut off. The 5 linguistic epochs are as follows: (1) add noun of the object, (2) add verb corresponding to the default action (push object A or pull object B), (3) add verb for opposite action (pull object A or push object B), (4) add both noun and default verb, and (5) add both noun and opposite verb. The various epochs are experienced by an organism in a random sequence. The same fitness formula is used as in the No-language case except that in the epochs when the opposite verb is used, the organism's action must reflect what the verb says, not what the object type would suggest by default.

In the third experimental condition, "Early-Language", organisms are exposed to all 11 epochs from the beginning of the simulation, i.e., from the first generation. For each condition, 20 replications of the simulations were run.

## Results and Discussion

The average performance of the organism in the three simulations is reported in Figure 3. For the two linguistic conditions, only the curve of the performance in the epoch with no linguistic input is reported, to allow a direct comparison among the three conditions. The No-language fitness curve grows until it stabilizes at around 15.8 successfully completed epochs. In the

Late-Language condition, at generation 1001 the population goes through a significant drop in performance. This appears to be due to the fact that the linguistic input reaches the hidden units through random weights that disturb the previous good performance. However, the behavior gradually improves and from around generation 1400 Late-Language organisms outperform No-Language organisms. The final number of successful tasks is 16.6 for the Late-Language condition. In contrast with this, the performance of the Early-Language population is less good than that of both the No-Language and the Late-Language populations (14.4).

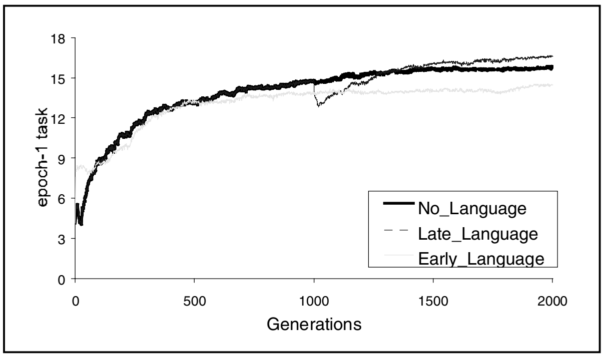


Figure 3 – Performance in epoch 1 (task without linguistic input) in the three experimental conditions

These results suggest an interesting hypothesis on language evolution and the interaction between linguistic and cognitive skills. To be adaptive language must be introduced at a later stage, after the cognitive abilities upon which it will be grounded have fully evolved. In this condition language has a beneficial influence on nonlinguistic behavior. If the evolutionary scenario involves both the practical task of pushing or pulling objects and the processing of linguistic signals from the beginning, it is more difficult to evolve populations with optimal performance in the practical task. Notice that if language is introduced later so that it can exploit the already existing (nonlinguistic) cognitive skills, the beneficial effects of language on nonlinguistic performance are observed not only when language is used together with nonlinguistic input (the language epochs) but also when there is no language and the organism is responding only to nonlinguistic input.

We will now focus on the Late-Language simulation to better understand why language has beneficial effects on nonlinguistic behavior and to analyze the differences between the two different classes of linguistic signals: nouns and verbs.

The 11 epochs of the Late-Language simulation can be grouped into 4 categories: (1) No-language, (2) Noun-only (the 2 epochs with and without retina input), (3) Verb-only (the four epochs with/without retina and with default/opposite verbs), and (4) Verb+Noun (the four epochs with/without retina and with default/opposite verbs).

Figure 4 shows the average performance for the three linguistic categories (categories 2-4) from generation 1000 to generation 1300. In the early generations, right after language has been introduced (from generation 1000 to generation 1100) the organisms' performance in the Noun-only epochs is higher than that of Verb-only and of Noun+Verb. Organisms learn to use nouns earlier than verbs to benefit their nonlinguistic performance. However, 100 generations later the disadvantage of the verb epochs disappears. Indeed, the performance for Verb-only and Verb+Noun epochs becomes stably better than that of Noun-only epochs.

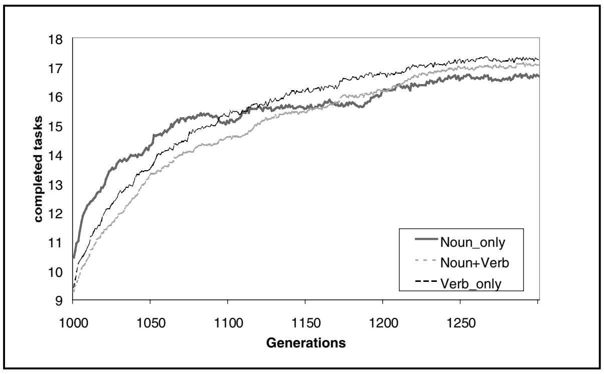


Figure 4 – Evolution of noun and verb use in the Late-Language simulation

The earlier advantage of nouns vs verbs can be explained by the fact that in the Noun-only epochs the task is consistent with what has been already learned without language up to generation 1000. Given this consistency with prelinguistic experience, nouns are easier to learn and they can benefit nonlinguistic performance earlier than verbs. On the contrary, with verbs organisms must learn to ignore some of the previously learned knowledge. When an opposite verb asks the organism to produce a new behavior (e.g., pull object A instead of pushing it, as previously learned) this is initially difficult to learn. Therefore, verbs can acquire an adaptive advantage only in later stages of language acquisition, when noun use has reached a good level of performance and stabilization and the individual can understand the more flexible nature of verbs, which can typically be predicated of a variety of arguments. This hypothesis could also explain the different stages of acquisition of nouns and verbs in children (Tomasello & Brooks, 1999). Verbs need a stable class of nouns to fully develop the potential and flexibility of their predicate-argument structure.

The Late-Language simulation can also be used to look at some aspects of the neural processing of language. To this purpose we analyzed the activation patterns in the second layer of hidden units (Figure 2), where sensory (retina+proprioception) and linguistic information come together and they both can have a role in determining the organism's motor behavior encoded in the output units. We used the activation patterns observed in these hidden units in the first cycle of each of the 18 motor tasks (9 for object A and 9 for object B). Each activation pattern can be represented as a point in the activation hyperspace of the hidden layer, with the 9 points corresponding to object A making up a "cloud" of points and the 9 points of object B making up another "cloud". We measured both the Euclidean distance between the centers of the two clouds and the size of each cloud as the average distance of the 9 points from the center of the cloud. (The points corresponding to objects/positions incorrectly handled were excluded from these calculations. On average, only 0.25 objects per epoch were misclassified.) The idea is that the successful accomplishment of the task requires that the different input patterns corresponding to the same object in different positions be internally represented as similarly as possible (small clouds) while the input patterns corresponding to the two different objects be represented as differently as possible (great distance between the two clouds).

The between-cloud distances and the sizes of the two clouds were computed for all 11 epochs. Then the data were averaged over the 4 categories of epochs: No-Language, Noun-only, Verb-only, and Noun+Verb. Figure 5 reports the average withn- and between-cloud distances at generation 2000. The between-cloud distances show a progressive increase from the No-language to the linguistic conditions. In an ANOVA test, these differences are statistically significant, except between the pair Verb-Only and Noun+Verb. A similar, but inverted, pattern of results is found for cloud size. The average size of a cloud decreases from the No-language to the linguistic conditions.

That language optimizes the representation of categories (i.e. increasing between-category distances and decreasing within-category sizes) has already been shown in other models (Cangelosi & Harnad, in press). What this model shows for the first time is that there are significant differences also between the three linguistic conditions, in particular between nouns and verbs. When the network is processing verbs, the size and distance of clouds is even better than when it is processing nouns.

How can we explain that verbs have even greater beneficial effects on nonverbal behavior than nouns? As we have shown, the beneficial effect of linguistic signals on nonlinguistic performance is due to the fact that linguistic signals induce better internal representations of reality. In our model, reality is internally represented in the neural network as the activation patterns observed in the higher layer of hidden units. The addition of language increases the distance between the two clouds of points (activation patterns) representing the two objects and decreases the size of the two clouds of points each representing one object. The language-modified clouds make it easier for the organism to select the appropriate action in response to the input. However, what is critical in internally representing reality is not to faithfully reflect the properties of the input but rather to prepare the motor output with which the organism must respond to the input. If the organism must be able to respond to the same object in different occasions with two different actions (push or pull) verbs are better than nouns in shaping the internal representations because while nouns covary with objects verbs covary with actions.

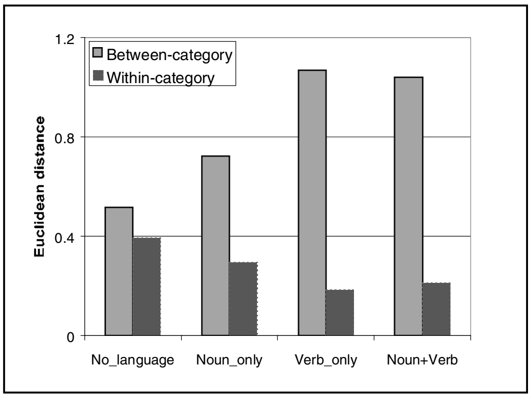


Figure 5 – Inter- and intra-categorical distances for the hidden representations at generation 2000.

## Conclusion

The present model focuses on the evolution of an innate language understanding ability for a language made up of nouns and verbs. Notwithstanding its obvious limitations, the models sheds some light on the reciprocal influences between language and nonlinguistic cognition, on the differences between nouns and verbs, and on the internal organization of neural networks that use language in an ecological context. Language has a beneficial effect on nonlinguistic cognition if it emerges on already existing basis of nonlinguistic skills, but not if it evolves together with them. The basis for this beneficial influence of language on behavior appears to be that language produces better internal representations of reality. That is, more similar representations of different situations that must be responded to with the same action, and more different internal representations of similar situations that must be responded to with different behaviors. Furthermore, verbs have a more beneficial effect on behavior than nouns because verbs, by their nature, tend to covary with the organism's actions while nouns tend to covary with the objects of reality that may be responded to with different actions in different occasions.

In this paper we have also done some comparisons between the computational model of language evolution and the literature on children's language acquisition and on neural processing of verbs and nouns. We are currently working on an extension of the object manipulation model to understand better the relations between language processing and sensorimotor knowledge (Martin et al, 1995). All in all, we believe this is a fruitful approach to the investigation of various adaptive, behavioral, and neural factors involved in the origin and evolution of language.

## Acknowledgments

Angelo Cangelosi's work for this paper was partially funded by an UK EPSRC Grant (GR/N01118).

## References

Batali, J. (1994). Innate biases and critical periods: Combining evolution and learning in the acquisition of syntax. In R. Brooks & P. Maes (Eds), *Artificial Life IV* (pp. 160-171), Cambridge, MA: MIT Press.

Cangelosi, A. (in press). Evolution of communication and language using signals, symbols, and words. *IEEE Transactions on Evolutionary Computation*.

Cangelosi, A. (1999). Modeling the evolution of communication: From stimulus associations to grounded symbolic associations. In D. Floreano et al. (Eds.), *Proceedings of ECAL99 European Conference on Artificial Life* (pp. 654-663), Berlin: Springer-Verlag.

Cangelosi, A., & Harnad, S. (in press). The adaptive advantage of symbolic theft over sensorimotor toil: Grounding language in perceptual categories. *Evolution of Communication*.

Cangelosi, A. & Parisi, D. (in press). *Simulating the Evolution of Language*. London: Springer-Verlag.

Cangelosi, A, & Parisi, D. (1998). The emergence of a language in an evolving population of neural networks. *Connection Science*, 10, 83-97.

Caramazza, A., & Hillis, A.E. (1991). Lexical organization of nouns and verbs in the brain. *Nature*, 349, 788-900.

Deacon, T.W. (1997). *The Symbolic Species: The Coevolution of Language and Human Brain*, London: Penguin.

Dessalles, J., & Ghadakpour, L. (Eds.) (2000). *Proceedings of the 3rd International Conference on the Evolution of Language*. Paris: ENST Press.

Di Ferdinando, A., Calabretta, R., & Parisi, D. (2001). Evolving modular architectures for neural networks. In R. French & J. Sougné (Eds.), *Proceedings of the Sixth Neural Computation and Psychology Workshop Evolution, Learning, and Development*, London: Springer Verlag.

Ellefson, M.R. & Christiansen, M.H. (2000). Subjacency constraints without universal grammar: Evidence from artificial language learning and connectionist modeling. In *The Proceedings of the 22nd Annual Conference of the Cognitive Science Society* (pp. 645-650). Mahwah, NJ: Lawrence Erlbaum.

Kaplan, F. (2000). Talking AIBO: First experimentation of verbal interactions with an autonomous four-legged robot. In A. Nijholt, D. Heylen, & K. Jokinen (Eds.). *Learning to Behave: Interacting agents. CELE-TWENTE Workshop on Language Technology* (57-63).

Kirby, S. (1999). Syntax out of learning: The cultural evolution of structured communication in a population of induction algorithms. In D. Floreano et al. (Eds.), *Proceedings of ECAL99 European Conference on Artificial Life* (pp. 694-703), Berlin: Springer-Verlag.

Martin, A., Haxby, J.V., Lalonde, F.M., Wiggs, C.L., & Ungerleider, L.G. (1995). Discrete cortical regions associated with knowledge of color and knowledge of action. *Science*, 270, 102-105.

Parisi, D. (1997). An Artificial Life approach to language. *Mind and Language*, 59, 121-146.

Rizzolatti, G. & Arbib, M. (1998). Language within our grasp. *Trends in Neuroscience*, 21, 188-194.

Schlesinger, M., & Barto, A. (1999). Optimal control methods for simulating the perception of causality in young infants. In M. Hahn & S.C. Stoness (Eds.), *Proceedings of the Twenty First Annual Conference of the Cognitive Science Society* (pp. 625-630). New Jersey: Lawrence Erlbaum.

Steels, L. (1997) The synthetic modeling of language origins. *Evolution of Communication*, 1, 1-37.

Steels, L. (2000) Mirror neurons and the action theory of language origins. *Proceedings of Architectures of the Mind, Architectures of the Brain*.

Steels, L., & Kaplan, F., (1999). Collective learning and semiotic dynamics. In D. Floreano et al. (Eds.), *Proceedings of ECAL99 European Conference on Artificial Life* (pp. 679-688), Berlin: Springer-Verlag.

Steels, L., & Vogt, P. (1997). Grounding adaptive language games in robotic agents. In P. Husband & I. Harvey (Eds). *Proceedings of the Fourth European Conference on Artificial Life* (pp. 474-482), London: MIT Press.

Tomasello, M., & Brook, P.J. (1999). Early syntactic development: A Construction Grammar approach. In M. Barrett (Ed.), *The Development of Language* (161-190). Philadelphia, PA: Psychology Press

Wiles, L., & Hallinan, J.S. (in press). Evolutionary computation and cognitive science: Modeling evolution and evolving models. *IEEE Transactions on Evolutionary Computation*, Special Issue on Evolutionary Computation and Cognitive Science.