The Emergence of Grammar in Communicating Autonomous Robotic Agents

Luc Steels
VUB AI Lab - Brussels
Sony Computer Science Laboratory - Paris
steels@arti.vub.ac.be

Abstract. Over the past five years, the topic of the origins of language is gaining prominence as one of the big unresolved questions of cognitive science. Artificial Intelligence can make a major contribution to this problem by working out precise, testable models using grounded robotic agents which interact with a real world environment and communicate among themselves or with humans about this environment. A potential side effect op this basic research are new technologies for man-machine interaction based on the negotiation of shared conventions.

1 Introduction

Artificial Intelligence researchers have for a long time been interested in the question whether and how machines can be built that are capable of natural language understanding and production (Winograd, 1972). This will only be possible if we go beyond statistical processing of language on which the success of present-day language technology relies. We need compositional analysis at a fine level of detail for parsing and production, and a procedural semantics establishing a two-way relationship between sentences and the world.

There has been vast progress on both these topics in the past decade of AI research. Usually it is assumed however that a language is in a fixed stable which can be circumscribed and then programmed or learned by giving a series of examples until the language state has been acquired. But human natural languages are constantly on the move. New words and phrasings appear all the time, new conceptualisations are continuously invented, and existing meanings shift often in subtle ways. This suggests that complementary to our efforts to understand the frozen state of a language, we need to understand the processes by which *new* language conventions and new conceptualisations of the world are created, negotiated and adapted by a community of users. This position is in line with the so-called bottom-up approach to artificial intelligence (Steels and Brooks, 1996) which advocates the construction of intelligent systems by evolutionary and adaptive techniques, because environments and tasks of agents will always be open-ended. Focusing on the processes that cause language conventions to originate and evolve may furthermore help us to understand the fascinating question how language itself might ever have emerged. This topic is receiving increasing attention lately in many fields interested in language, and computational modelling is playing an important role to help formulate and test concrete hypotheses (see overviews of this research trend in Steels (1997) and Hurford, et.al. (1999)).

2 The origins of word meaning

Our own work has focused in a first phase on the problem how individual words, or groups of words without grammar, might be associated with meanings by a distributed negotiation process in a population of autonomous grounded agents. Since language is about the real world as experienced through sensors and actuators, we felt a strong need to experiment with physical robots. We have built and experimented with several kinds of robotic bodies, ranging from small mobile robots based on Lego technology (Steels and Vogt,1997) to fixed steerable cameras (Steels,1997), and more recently animal-like or humanoid shaped robots. The steerable cameras were used in our most important experiment to date, the Talking Heads experiment (see figure 1).

Language evolution clearly takes place in populations, so we needed the ability to do experiments with large numbers of agents. This was achieved by allowing agents to be loaded in different bodies and transporting their mental states from one body to another so that agents can engage in interactions from many perspectives and in many different environments without requiring that they

Figure 1: The 'Talking Heads' experimental setup with two steerable cameras capturing images of geometrical figures in front of them. Each camera is used by an agent in a grounded language game.

physically move around. We have built a 'cognitive teleporting' infrastructure such that physical robots can be networked through the Internet and thus 'receive' agents. The use of Internet has the additional advantage that experimenters can create agents through the web and send them around to different places and that experiments can be monitored and inspected from wherever the experimenter happens to find herself. It also has enabled us to set up large-scale experiments involving human-agent interaction.

This general infrastructure has been fully operational now for over a year and experiments have been done with populations of up to 4000 agents over a period of several months. Robotic sites capable to receive agents have been active in Paris, Brussels, Amsterdam, London, Tokyo, Lausanne, and Antwerp. In total, close to a million grounded situated language interactions have taken place and agents have made tens of thousands of travels over the Internet between different sites. We hope that this infrastructure will become a general test ground for exploring various theories on the emergence and evolution of language understanding and production in autonomous situated agents. Indeed, the infrastructure is neutral with respect to which theory of language one adopts and is therefore ideally suited to compare the adequacy and performance of different theories.

More concretely, the environment of the Talking Heads experiment consists of geometric coloured figures pasted on a white board. The agents have a limited set of sensory channels (aspects of color, shape, and position), mechanisms for segmentation, for identifying the most salient features of the objects in the scenes before them, and for categorising objects based on evolving discrimination networks. The agents have a lexical component based on a 2-way associative memory associating words with meanings and meanings with words. They play a game called the guessing game in which one agent tries to identify an object in the scene captured by a camera to another agent using verbal means. When the agents do not have sufficient categories or sufficient words, the discrimination trees or lexical memory expands.

In our experiments so far, we have seen that shared lexical systems evolve given relatively stable environmental circumstances and as long as the population flux is not too high (Steels and Kaplan, 1998). The reason for this self-organisation is a positive feedback loop between the use of a word-meaning pair and the success of that pair: Speakers and hearers prefer the word-meaning pair with the highest score and update their score based on success in the game. We have also seen that the lexicon keeps evolving due to fluxes of agents, new environments, perturbations, etc. (Steels and Kaplan, 1999b).

The remainder of this paper focuses on experiments that go significantly beyond these early results because they address the problem of the origins and acquisition of grammar. This implies that two problems need to be attacked: the origins of more complex compositional meaning and the origins of grammatical conventions to express or parse such complex meaning.

3 Two approaches to grammar

Clearly natural languages exhibit universal tendencies (Greenberg, 1966). For example, if a language shows a distinction between subject, object(s) and verb, then we often see that the subject comes first followed by the verb and then the remaining object(s), as in English: "John likes novels", or French "Jean aimes les romans". However not all languages follow this SVO pattern, for example, some like Japanese use SOV. Of course when a language does not employ a clear distinction between subject, verb, and object(s), then talking about such patterns is largely irrelevant. For example, "Marlunnik ammassattorpunga" (Greenlandic Eskimo) translates literally as "Two-instrument-plural sardine-eat-1st-singular-Indicative" to mean "I ate two sardines". The object has been incorporated into the verb and case dependencies and other information is expressed using lexical morphemes functioning as affixes (Van Valin and LaPolla, 1997). This suggests that although there are universal tendencies in natural languages, these tendencies should not be construed as immutable general laws.

3.1 Universal Grammar

The first possible explanation for the universal tendencies is that they form part of the human genetic make-up, similar to the way we all have two hands and walk upright. Thus the different grammatical categories (like adjective or subject), the general patterns on how members of a category can be combined, the types of meanings that can be expressed, parsing and production strategies, language acquisition strategies, the mapping from syntactic form to meaning and back, the ability to engage in dialog, and so on, are in this view all innately known. In such a scenario, language would be largely transmitted genetically and so we can expect there to be a section of the brain especially wired for natural language. The obvious differences between languages and the evolution of language through time can be explained by assuming that there are *parameters* in the universal 'Bauplan'. These parameters are set by the brain during development under cultural influence.

This explanation was first put forward and defended by Chomsky and linguists in the formal generative tradition (Chomsky,1981) (Lightfoot,1991). For the present purpose, namely building a language understanding - production system capable to acquire and adapt to the language in its environment, the approach prescribes that we should implement the Universal Language Acquisition Device and show how it gets instantiated by exposure to a specific language. Most computational linguists implicitly follow this assumption because they build in the complete machinery for handling language. Also the evolution of language or the emergence and expansion of a language can in principle be computationally modelled from this perspective. Evolution is assumed to take place when parameter settings shift, and expansion is guided by instantiating choices available in universal grammar. Concrete computational experiments in this direction are discussed in Briscoe (1999).

3.2 Complex Adaptive Systems

The Universal Grammar approach has been so popular during the second half of the 20th century in linguistics that we would almost forget that there is an alternative, which we can call (to use modern terminology) the complex adaptive systems approach. Language communication is now seen as a continually changing complex adaptive system, not unlike an economy or an evolving ecosystem. Individual language users are the units in the system and they engage in local interactions with each other. Global sharing arises through self-organisation and the structural coupling between the evolving language competence and the processes developing meaning. Each language user has a specific state of knowledge about their language, which they use for their own communicative behavior.

From the viewpoint of the complex adaptive systems approach, a language system is adaptive in two senses: Individuals produce and interpret sentences partly in a routinised way (otherwise we cannot explain how they can speak so fast). But occasionally the language needs to be expanded by a speaker to deal with novel concepts and situations and these expansions need to be learned so that they can propagate to the rest of the population. Language users optimise their behavior to be more successful in future communications and to minimise the energy and memory resources they need to apply. Because individuals adapt their grammars, the language as a whole changes and evolves and this in turn determines how individuals must change if they want to be understood by the rest of the population. Systematicity is always temporary and perhaps less pronounced than a Universal Grammar approach tends to suggest (Hopper, 1987).

The universal tendencies we observe in language are now considered to be emergent properties, because the language evolves in a selectionist fashion under various constraints. The constraints come from the tasks the language system has to satisfy, namely communication about the world between autonomous agents, from the sensori-motor apparatus and cognitive architecture available to human language users to achieve this task, from the demands of real world communication without telepathy, with limited time, and with a noisy transmission medium, and from the learnability requirement. Instead of a genetic origin of language, the complex adaptive systems view leans towards a cultural origin, transmission and evolution of language.

This alternative view resonates strongly with recent work in complex systems in biology and economics, but has also been defended within linguistics itself. For example, several researchers have tried to explain why natural languages show specific distributions in their sound repertoires based on functional, cognitive and sensori-motor constraints (see e.g. Lindblom, MacNeilage,Studdert-Kennedy, 1984). In linguistics, this view is most compatible with the work of linguists taking an empirical attitude, looking how new constructs like adjectives, articles or auxiliaries may appear or disappear in a language or how languages may gradually shift from a morphological strategy for expressing case to a word order strategy (Heine, Claudi, Huennemeyer, 1991).

A complex adaptive systems suggests a completely different way to evolve language understanding/producing systems, compared to the one pursuing a Universal Grammar approach. We should now build a general cognitive architecture and a sensori-motor apparatus that approaches that of humans, and then perform experiments to see whether such artificial systems can develop languages with the complexity and characteristics of human languages and whether they are able to acquire natural language within similar time and data constraints as experienced by humans. The Talking Heads experiment discussed earlier follows this line of research for studying the origins of word meaning. Examples of such work for grammar are given in Batali (1999), Kirby (1999), or Steels (1997). The remainder of this paper reports more deepply on work along these lines.

4 The origins of compositional meaning

Given the size and complexity of natural language meaning, it is obviously a very deep challenge to find mechanisms that can explain how meaning emerges in interactions with the environment and with other language users. We start from the assumption that at least in its primary function language is intended for communication. Communication is a form of coordinated action. The speaker hints at an action or a series of actions that she wants the hearer to perform. These actions are either physical actions in the world ("Give me that book") or mental actions to focus on certain items in the context ("the book on top of the refrigerator in the kitchen"), to store facts for later use ("the book is no longer there"), etc. To perform a communication, the speaker therefore has to plan what actions she wants the hearer to perform. A possible plan is then translated into an utterance satisfying the conventions adopted by the language community and this plan needs to be decoded and then interpreted by the hearer. Success in communication arises when the effect of the hearer's actions is the one desired by the speaker. Part of the difficulty of natural language communication is that the plans formulated by the speaker are only vaguely hinted at. A lot of intelligence is required from the hearer to interpret them correctly. Because we want to integrate verbal and non-verbal communication, some of the actions may also be actions in the world like moving the head or performing a pointing gesture.

It follows that the construction of complex compositional meaning can be compared to a planning problem and all the techniques developed in AI to tackle this problem become relevant. Specifically, we have used techniques from constraint propagation for the interpretation of plans, we use search in the space of possible plans to find a plan capable to satisfy the current communicative goal, and chunking to abstract a successful plan into a new component for future usage. Because of chunking, the agent gradually builds up a library of directly-usable complex components. There is not enough space here to go into great technical detail, nor to summarise the main technical precursors of our implementation. The remainder of this section just gives a very sketchy idea.

The first unit of the system is an *object store* which contains the objects (based on segments of the image), features of the objects (still in a continuous space scaled between 0.0 and 1.0), and aggregations such as sets or sequences of objects, which are all relevant for the scene as captured by the camera. The object store is initially filled by a battery of standard low level sensori-motor processing routines that segment the image, detect a variety of features and reflect motor states.

The second unit of the system is a *component store* which contains a set of components, the building blocks available for the planning process. Each component can be viewed as a constraint logic program capable to satisfy a particular goal, like filtering a set of objects into two subsets depending on whether or not they satisfy a category, ordering a sequence of objects based on a category, etc. Each component usually maintains also a *knowledge structure* which contains the evolving knowledge needed for the operation of the component. Many of these components are similar to the operations used in Montague-style semantics. They are augmented with operations that use neural-network style approaches, such as a Kohonen network to categorise n-dimensional feature vectors into distinct classes.

Here is an example component, which we call COMPARE-PROTOTYPE. It performs a filtering operation by comparing the elements of a set (further called the source-set) based on their distance to a prototype. The set of prototypes available to COMPARE-PROTOTYPE start out on the basis of concrete input examples of which most of the contingent properties are gradually stripped away. COMPARE-PROTOTYPE retains those elements of the set that are close to the prototype and collects them in another set (further called the object-set). Such a network appears useful for the procedural semantics of many nouns. The set of prototypes is in this case equal to the knowledge structure associated with this component. Another component, which we call COMPARE-AVERAGE, performs a similar operation but now by comparing the values of the elements along a particular dimension (e.g. Horizontal Position) to their average in the present situation, retaining those that are less than or larger than the average. Such a primitive network is relevant for the procedural semantics of concepts like 'left' or 'right'. The relevant knowledge structure in this case maintains a possibly growing list of the possible dimensions and possible comparisions.

Apart from primitive components, there are also more complex components which consist of an assembly of component instances put together by equating their slots, thus forming a constraint network. Thus for a phrase like "the left table" as in the sentence "put the box on the left table", a complex combination is required, like the one below:

```
(IDENTIFY-OBJECT-WITH-PROTOTYPE-AND-OPERATOR
      Object Prototype Operator) :=
  (EQUAL-TO-CONTEXT Object-set)
  (COMPARE-PROTOTYPE Object-set-2
        Object-set Prototype)
  (COMPARE-AVERAGE Object-set-3
        Object-set-2 Operator)
  (UNIQUE-MEMBER Object Object-set-3)
```

with Prototype bound to the [table] prototype and Operator to [< horizontal-position]. The EQUAL-TO-CONTEXT component maintains the object-set equal to the elements in the present context. UNIQUE-MEMBER picks out one element from an object-set (here Object-set-3) which is assumed to be a singleton. The various arguments (object, prototype, object-set, etc.) are slots that are filled in or used by each subcomponent.

A study of natural language quickly reveals that the computational processes required in the interpretation of semantic plans needs to be very non-trivial, which explains why we need constraint propagation as opposed to simple sequential control. First of all it will have to use data flow (rather than explicit control flow) in the sense that information should propagate in any direction whenever possible. For example, in the phrase "the ball rolls to the edge of the table", we can only uniquely identify the ball after we have identified the edge and also the table. But there may be more than one table, there is any case more than one edge for each table, and there might be several balls - but perhaps only one rolling to the edge of the table. So the computation must take the form of an attempt to find a set of fillers for all slots that is internally consistent and compatible with the present context.

Second, the computation will have to examine many different possibilities at the same time, as already illustrated by the previous example. In the current implementation this is done by exploring in parallel many different possible worlds, expanding into competing worlds when there are many different hypotheses and collapsing worlds in which some relations are invalid, i.e. in which the underlying network signals a failure to establish the relation. This technique has again been adopted from recent research in constraint propagation.

Third, the computation needs to be able to proceed in any direction. Consider the component IDENTIFY-OBJECT-WITH-PROTOTYPE-AND-OPERATOR discussed above which uses a prototype and an operator. While performing the planning process for speaking, fillers for the Prototype and Operator need to be found given bindings for Object-set-2 and Object-set-3, so the component acts as a generator of possibilities. Conversely, during understanding the Prototype and Operator are given and Object-set-2 and Object-set-3 need to be found. This points again in the direction of constraint propagation methods as the fundamental basis for natural language semantics.

Components need to be connected to goals. The speaker starts from a communicative goal that needs to be achieved. The planning process has been organised as a search in the space of possible combinations of primitive components. Rather than reasoning from first principles based on pre- and postconditions as in traditional logic-based planners, we have implemented a system where a combination of components is assembled and simply tried. A successful combination is abstracted out as a chunk and associated with the communicative goal that it managed to achieve. Examples of chunks for identifying objects might be: One that uses a prototype, like in "the box", one that uses a prototype and an operator, like in "the small box", one that uses a relation, for example to identify a location as in "left of the ball", etc. By the chunking process a repertoire of complex components with associated goals gradually arises and so planning becomes mostly the retrieval of high level ready-made plans rather than the microplanning from scratch. The repertoire of stereotyped plans derived from chunking is organised in a hierarchy based on the goals they achieve and their developmental history. The structure of this hierarchy can be exploited for searching through the space of applicable components while planning. It appears also as the backbone for the grammar. The component repertoire that typically emerges point already to some universal tendencies in grammar. For example, the distinction between noun phrases and verb phrases reflects a distinction between the major communicative goals: identify-object and describe-situation. Because components have slots, they naturally lead to the evolution of case systems expressing what items fill what slots.

5 The emergence of grammar

Before we can attempt to make a computational model of the origins of grammar, we must reflect on the question what grammars are for. Several hypotheses have been put forward on this matter. Specifically, Kirby (1999) and Batali (1999) have proposed and shown in computer simulations that grammar emerges because languages need to pass through a learning bottleneck. A language, in order to survive, has to be learnable by the next generation. This is easier when a language exhibits structural regularities. These structural regularities automatically arise by over-interpretation or re-use of existing structures. So agents that use memory-based language processing with modest forms of abstraction (e.g. analogy) automatically generate a form of grammar, where grammar is here seen as structural regularities in the language.

Although there is a lot to say for this idea, I propose instead the hypothesis that the lexicon expresses the basic semantic items in the semantic plans exchanged between the agents but that grammar is required to hint at what kind of plan is intended, in other words what should be done with the basic items. I adopt therefore a functional and cognitive stance towards language as opposed to a structuralist one, i.e. form is not arbitrary but has a communicative or semantic function. For example, the same concept [large] (which involves the area dimension and a greater than comparison) can be used in multiple ways: "the large box" (large with respect to the other objects in the context), "the largest box" (the one with largest size of all objects), "the box enlarges" (the size becomes larger), "the box larger than the ball" (comparison with respect to size of another object). Each usage involves the same basic lexical item ("large") but with various grammatical form elements added. Because the speaker assumes that the hearer has a similar planning apparatus and plan repertoire, grammatical expressions are only needed when it is not clear what plan is intended. But, as the agents' repertoire of possible plans becomes larger, they must seek ways to clarify to the hearer which plans should be invoked.

Memory-based processing still plays a key role however, partly for reasons of efficiency. Ready-made language solutions must immediately be triggered with often-used semantic structures to explain the speed of language communication. And because memory-based processing is used, structures might be left in the language which have lost their original functional significance, but which are nevertheless preserved by example-based transmission. The next question is then where the original form-meaning mappings come from if they do not come from the spontaneous re-use and over-interpretation of examples.

I hypothesise that language speakers use a variety of different strategies to map form to meaning, each strategy yielding a different sort of style of expression. An example strategy is to use lexical tags attached to basic lexical items to indicate the slots and components. Another strategy is to use word order for some of the same information. The first strategy would be used for example in a language with a strong case system expressed through affixes. English uses for the same problem a more syntactic approach where word order expresses case relations (i.e. which slot is filled by what). Languages use multiple strategies and strategies shift during the history of the language. Often there is a period of productive use of a strategy, followed by sedimentation and fossilisation, and the resulting debris is then used as material for other strategies. Different languages may nevertheless use the same strategies for some aspects of meaning, explaining why we see the universal tendencies discussed earlier. This still does not mean that they have to be innate. The language strategies should themselves emerge and be learned by the agents exercising their cognitive apparatus for the communication task.

A strategy always needs four components:

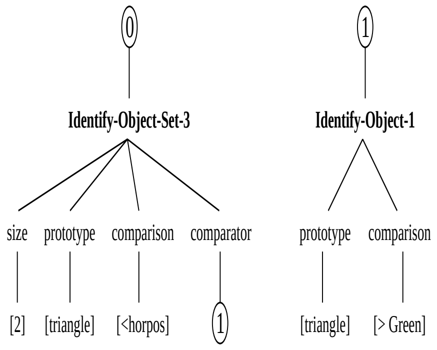

Figure 2: Example of semantic trees produced by the planning process. The trees reflect a meaning invoked by the phrase "The two squares left of the green triangle."

- A method for routine *production*, which requires specific storage structures like an associative memory for morphologically-oriented strategies or a pattern memory for syntactically-oriented strategies.
- A method for expanding the grammar and lexicon when routine solutions are not available but novel structures need to be verbalised. I call this the *invention* strategy.
- A method for routine *parsing*, which requires similar knowledge structures as for production but a different usage of the same knowledge.
- A method for *learning* unfamiliar constructions.

In any case, the structures produces by a strategy are always subject to chunking and memorisation for efficient memory-based retrieval.

In our implementation experiments so far, we use similar structures for syntax and semantics. The semantics are derived from the plans generated by the planning process described in the previous section and converted into trees, one for each component that is used in the network. Each tree has a unique index for crossreference in other trees, a component, its various slots, and the fillers of the slots which are either items or references to other trees. A syntactic structure is also seen as a plan, now for producing the expression by assembly of its component parts. This plan is recognised when parsing an expression. An example component in a syntactic structure is ORDERED-GROUP-OF-3 which has 3 slots to be filled by other groups or individual words. Words are also components which have slots for affixes.

A variety of strategies has already been implemented. Here is an example strategy which performs *full lexicalisation*, in other words all the aspects of a semantic structure have corresponding morphemes. Each path in each tree corresponds to a word and the order of morphemes in a word goes from top to bottom. This strategy is of course artificial and not used as such by any natural language, but by researching such strategies we can gain progressively a better insight into those that are more realistic and clearly used in natural languages and those that are not.

The production and invention component of a full lexicalisation strategy is straightforward. The agent just looks up in his lexical memory what the best lexicalisation is for each node in the tree (i.e. the one with the highest score), traverses each tree from top to bottom and thus collects the different words. If on the way there is a node which has no expression in the language yet, then a new morpheme is created. Note that the index of a tree is also lexicalised and then re-used similar to the use of pronouns. Each association has a score within the memory of the agent, reflecting how strongly this association is believed to be valid in the language according to the feedback on language interaction this agent has received.

The parsing strategy is less trivial because of the unavoidable ambiguity and uncertainty causing one word to have more than one possible meaning (and one meaning more than one possible world). But it can be done in the following way: The hearer collects all the possible meanings of each morpheme from the lexicon. Then the hearer collects all possible uses of each meaning in terms of the meaning repertoire. Indeed, the hearer can derive for each bottom-node in the tree what the possible components and slots are in which it fits. It can derive for each possible slots in which components it can be. This generates a lot of possible hypotheses which are then shaken out to retain only those that are internally consistent. The hearer then tries to interpret the remaining semantic structures in terms of the situation of the communication and ideally arrives at one possible plausible result.

When the communication fails, the learning strategy must come in action. The one used in the present implementation of the full lexicalisation strategy works as follows. The hearer collects the best possible interpretation of the utterance transmitted by the speaker. In other words, the hearer performs his own conceptualisation of the communicative goal and thus guesses what kind of semantic structure the speaker might have intended. Usually there are many possible partial matches between the utterance and the target semantic structure. The best one is chosen and then the language memory is updated. New associations might be stored, the score of associations that are part of the successful match are increased and its competitors decreased. Also in a successful communication, speaker and hearer must adapt their scores to be more successful in future games.

Simulations with robotic agents have shown that this full lexicalisation strategy is effective in the sense that agents build up a shared repertoire of morphemes to express their expanding repertoire of stereotyped plans. Word order is entirely irrelevant in this language. The order of morphemes was adopted under pressure from the difficulty of language acquisition, otherwise there would be an explosion of possibilities. This shows that we always have to keep the four aspects of language processing in mind: parsing, production, invention *and* learning.

Another implemented strategy is to use word order for expressing which slots are present, as opposed to lexical tags. This can be achieved by associating with each component a pattern that prescribes the order in which the components need to be presented. Yet another strategy is to be much more economical in expression. The speaker starts by expressing the bottom nodes of the semantic structures and only includes higher level nodes (through word order or lexical tags) when absolutely required. The speaker can simulate the difficulty of parsing by first parsing the utterance internally, and only adds more information when the utterance is ambiguous with respect to the semantic information and the state of the grammar. This strategy results in a "telegraphic style" of the language, reminiscent of pidgin or two year old "protolanguages", particularly in the early phases when not many ambiguities exist yet due to the limited size of the agents' meaning repertoire.

6 Conclusion

Building natural language understanding and production systems requires that we come to grips with the problem how new language conventions may arise in a distributed population of grounded agents. We have opted for a complex adaptive systems approach to the emergence of grammar. Agents come equiped with a basic cognitive apparatus but the specific meanings and language conventions arise by a negotiated process grounded and situated in interactions about the world. Some first exploratory simulations have shown the viability of the approach although much more basic research is needed to complexify both the interaction with the world, the repertoires of primitive components, the communicative goals, and the language strategies agents use.

7 Acknowledgement

This research was conducted at the Sony Computer Science Laboratory Paris with major contributions by Angus McIntyre and Frederic Kaplan, and at the VUB AI laboratory with specific contributions from Joris van Looveren. The VUB research was financed by a GOA grant from the Flemish government.

8 References

Batali, J. (1998) Computational Simulations of the Emergence of Grammar. pp. 405-426 in: Hurford, et.al. (1998)

Briscoe, T. (1999) Grammatical Acquisition and Linguistic Selection. In: Briscoe, T. (ed.) (1999) *Linguistic evolution through language acquisition: formal and computational models.* Cambridge Univ. Press, Cambridge.

Chomsky, N. (1981) *Lectures on government and binding.* Foris Publications, Dordrecht.

Greenberg, J. (1966) Language universals. Mouton, Den Haag.

Heine, B., U. Claudi and F. Huennemeyer (1991) Grammaticalisation. A conceptual framework. The University of Chicago Press, Chicago.

Hopper, P. (1987) Emergent Grammar. Berkeley Linguistic Society 13, 139-157.

Hurford, J, C. Knight and M. Studdert-Kennedy (eds.) (1999) *Approaches to the Evolution of Human Language.* Cambridge Univ. Press. Cambridge.

Kirby, S. (1999) Function, Selection and Innateness: the Emergence of Language Universals. Oxford University Press, Oxford.

Lightfoot, D. (1991) How to set parameters: arguments from language change. The MIT Press, Cambridge Ma.

Lindblom, B., P. MacNeilage, and M. Studdert-Kennedy (1984) Self-organizing processes and the explanation of phonological universals. In: Butterworth, G., B. Comrie and O. Dahl (eds.) (1984) Explanations for Language Universals. Walter de Gruyter, Berlin. pp. 181-203.

Rice, M. (1996) Toward a Genetics of Language. Lawrence Erlbaum Ass. New Haven.

Steels, L. (1997) The synthetic modeling of language origins. *Evolution of Communication*, 1(1):1-35.

Steels, L. and Kaplan, F. (1998) Situated Grounded Word Semantics. In *Proceedings of IJCAI-99*, Stockholm. Morgan Kauffman Publishing, Los Angeles. p. 862-867.

Steels, L. (1998) The Origins of Syntax in visually grounded robotic agents. Artificial Intelligence 103, 1-24.

Steels, L. and Kaplan, F. (1999) In *Proceedings of COLING-ACL*, Montreal, August 1998, p. 1243-1249.

Steels, L. and P. Vogt (1997) Grounding Adaptive Language Games in Robotic Agents. In Harvey, I. et.al. (eds.) *Proceedings of ECAL 97*, Brighton UK, July 1997. The MIT Press, Cambridge Ma.

Van Valin, R. and R. LaPolla. (1997) Syntax. Structure, meaning and function. Cambridge University Press, Cambridge.