Chapter 16
Modeling the Formation of Language in Embodied Agents: Conclusions and Future Research

Luc Steels and Vittorio Loreto

Abstract This chapter draws some conclusions from the computational and mathematical models of emergent symbolic communication systems reported in the earlier chapters. It also strongly pleads for a stronger interaction between linguistics and other human sciences studying similar issues.

1 Introduction

Clearly, huge progress has been made recently in building Embodied Communicating Agents that use symbolic (conventionalised) communication, which we briefly review here with respect to the four challenges discussed in Chap. 13 (i.e., language games, concept formation, lexicon, and grammar), plus the issue of embodiment (a detailed review of the actual experiments which permitted this progress has been done in Chap. 14).

2 Embodiment

Pioneering works on the emergence of language in embodied agents as carried out in the late nineties was mostly based on pan-tilt cameras (as in the 'Talking Heads' experiment). Recent research in the field (most of which have been reviewed in Chap. 14) has been based on more complex robots. Some experiments have been conducted with the AIBO autonomous dog-like robots, particularly in the domain of spatial language (Loetzsch et al. 2008b; Steels and Loetzsch 2008), and as humanoid robots become progressively available (such as the Sony QRIO, the Aldebaran NAO, and the IIT ICub), an increasing number of experiments targets these

L. Steels (✉)
Sony-CSL, Paris, France
e-mail: steels@arti.vub.ac.be

V. Loreto
"Sapienza" Università di Roma, Rome, Italy
e-mail: loreto@roma1.infn.it

S. Nolfi, M. Mirolli (eds.), *Evolution of Communication and Language in Embodied Agents*, 283
DOI 10.1007/978-3-642-01250-1_16, © Springer-Verlag Berlin Heidelberg 2010

platforms (such as the Grounded Naming Game; Steels et al. 2008). This scale-up of robotic complexity has never before been achieved in research in embodied communication and obviously required a dramatic scale-up in the sensori-motor systems used by the robots. They also provided an important scale-up in the complexity of what the agents could communicate about and thus in the opportunities for experimentation. Future work in this area will be driven forward by the requirements of increasingly complex experiments. This will involve a further scale-up in terms of the sensori-systems used by the robots, world complexity, robot-world interactions and the interactions among the robots themselves.

3 Language Games

In the research described above we significantly extended the state of the art from the naming and guessing games to several other games, such as an action game where robots execute actions for each other (Steels and Spranger 2008), a construction game where robots have to construct an object based on instructions from the other robot (Loetzsch et al. 2008a), a description game in which agents have to describe dynamic real-world scenes to each other (Steels 2004; Trijp 2008), etc. Important technical work has been done on implementing scripts on autonomous robots for playing language games (Loetzsch et al. 2008c), particularly for achieving joint attention (Steels and Loetzsch 2008).

Future work in the domain of language games will investigate which interaction scripts are required for other linguistic phenomena. This may involve a scale-up in terms of the number of agents involved in a single game (e.g. a speaker and multiple hearers) and turn-taking (i.e., longer discourse or dialogue structures). For example, some phenomena such as anaphora, information structure markers and determiners may require more lengthy discourse. As with all the other areas of investigation, however, the focus will lie on defining a language game in its most essential form in order to avoid interference from other communicative pressures.

4 Concept Formation

Pioneering works in this area used various concept formation mechanisms such as discrimination trees (Steels 1996), radial basis function networks (Steels and Belpaeme 2005), etc. Each experiment used its own mechanism, depending on the nature of the semantic domain. The research reviewed in Chap. 14 has further explored these concept formation mechanisms. However the big breakthrough has come from a new approach to concept formation which uses second order semantics (IRL; Steels and Bleys 2005). Agents now have a library of mechanisms for conceptualization and concept formation in the form of (procedural) constraints and they compose these mechanisms when they have to come up with a complex conceptualization. This allows us to deal with semantics of phrases like “the very big blue ball" which go beyond simple first order predicate calculus. The IRL system uses genetic programming methods to come up with a possible "plan" for communication and chunks found solutions so that they can be more easily found during later communicative interactions.

The main challenge for the future lies in the integration of the IRL system with work on embodiment on the one hand, and the language system (Fluid Construction Grammar) on the other. In the former case, a library of "cognitive primitives" needs to be collected, i.e., several concept formation mechanisms (such as discrimination trees) need to be implemented and represented through IRL constraints. With respect to the integration with FCG, we need to investigate how IRL networks relate to lexical entries and grammatical constructions, and we need to examine how agents can exploit IRL for solving communicative problems when there are not enough linguistic conventions at hand.

5 Lexicon

The Talking Heads experiment already contained solid ways to form a lexicon based on a lateral inhibition dynamics (Steels et al. 2002). In this respect, the work reviewed in Chap. 15 has given a formal foundation to the knowledge acquired through robotic experiments. Furthermore, the scaling laws, convergence properties, etc. have now been studied thoroughly from a complex systems point of view (Baronchelli et al. 2006; Vylder and Tuyls 2006). The generality of the approach has also been demonstrated further by new experiments in spatial language, body language, etc. Some alternative approaches have been studied to achieve more flexible lexicons (Wellens et al. 2008), that handle more easily the inevitable combinatorial explosions that come when the meaning of a word is not clear (the Gavagai problem).

With the firm basis that we established with respect to the dynamics of vocabularies, future work on lexicon formation should mainly focus on specific domains and exploit the lexicon as a stepping stone to richer embodiment and grammatical languages. For example, our work on spatial vocabularies (Steels and Loetzsch 2008) did not start from scratch but used all the insights gained from previous experiments, which allowed us to focus more thoroughly on the enormous challenges of grounding and embodiment. Likewise, our current understanding of lexicon formation enables us to tackle even more ambitious questions such as metaphorical extension (e.g., spatial expressions are very often extended to the temporal domain), polysemy (e.g., lexical drift and extension to multiple contexts) and grammar (i.e., how these spatial words become part of idiomatic expressions and grammatical constructions such as adverbial phrases and argument structure constructions about caused motion).

6 Grammar

Finally, since a few years ago, no significant experiments had been carried out yet in the domain of grammar, despite several attempts. It was only very recently (with the Case Experiment reviewed in Chap. 14) that clear breakthrough experiments have been achieved. This has required first of all a solid implementation of a more sophisticated formal and computational framework for language processing, the Fluid Construction Grammar framework (De Beule and Steels 2005; Steels and De Beule 2006). This framework uses a wide range of existing techniques from computational linguistics (feature structures, unification, search spaces, etc.) but adds novel new mechanisms, such as the J-operator for handling bi-directional hierarchies (De Beule and Steels 2005). The second step was to find the diagnostic and repair strategies that would allow for the emergence of grammar (Steels 2004). We have focused on case grammar, as this is widely seen as a core component of grammar and has been intensely studied in linguistics. A breakthrough experiment (the case grammar experiment) has now shown that a given set of diagnostic and repair strategies is indeed able to show the emergence of both the semantic roles (agent, patient, etc.) and the markings of these roles, in a way that is compatible with Construction Grammar approaches (Steels 2004; Trijp 2008).

These first breakthrough experiments show that all the previous insights gained from previous work on lexicon formation can be successfully moved to the domain of grammar. Future work has to tackle multiple grammatical domains (e.g., space, tense-aspect and event structure) and investigate more complex learning operators, innovation mechanisms and alignment strategies. An enormous challenge also lies in achieving a tighter integration of the experiments on grammar with the work on embodiment and semantics (IRL). Finally, research on grammar has to investigate how languages evolve over time, for example how a particular language can shed its case markers and evolve towards a word order-based grammar.

7 Mathematical Modeling

Statistical physics has proven to be a very fruitful framework to describe phenomena outside the realm of traditional physics (Loreto and Steels 2007). The last years have witnessed the attempt by physicists to study collective phenomena emerging from the interactions of individuals as elementary units in social structures (Castellano et al. 2009). These macroscopic phenomena naturally call for a statistical physics approach to social behavior, i.e., the attempt to understand regularities at large scale as collective effects of the interaction among single individuals, considered as relatively simple entities. This is the paradigm of the complex systems: an assembly of many interacting (and simple) units whose collective (i.e., large scale) behavior is not trivially deducible from the knowledge of the rules that govern their mutual interactions. This scenario is also true for problems related to the emergence of language. As linguists begin to get access to more and more data from systematic recordings and the massive volume of text appearing on the World Wide Web, and as they look at new language-like communication systems that have emerged recently—such as text messaging protocols for use with mobile phones or social tagging of resources available on the Web—doubts arise whether human communication systems can be captured within a static picture or in a clean formal calculus. The static picture is giving way to a view where language is undergoing constant change as speakers and hearers use all their available resources in creative ways to achieve their communicative goals. This is the point that looks at language as an adaptive evolving system where new words and grammatical constructions may be invented or acquired, new meanings may arise, the relation between language and meaning may shift (e.g., if a word adopts a new meaning), the relation between meanings and the world may shift (e.g., if new perceptually grounded categories are introduced). All these changes happen both at the level of the individual and at the group level, the focus being on the interactions among the individuals as well as on horizontal, i.e., peer to peer, communications. In this new perspective, complex systems science turns out to be a natural allied in the quest for the general mechanisms underlying the emergence of a shared set of conventions in a population of individuals. In this respect, statistical physics brings an important added value. In most situations qualitative (and even some quantitative) properties of large scale phenomena do not depend on the microscopic details of the process. Only higher level features, as symmetries, dimensionality or conservation laws, are relevant for the global behavior. With this concept of *universality* in mind one can then approach the modelization of social systems, trying to include only the simplest and most important properties of single individuals and looking for qualitative features exhibited by models.

A crucial step in this perspective is the comparison with empirical data which should be primarily intended as an investigation on whether the trends seen in real data are compatible with plausible microscopic modeling of the individuals, are self-consistent or require additional ingredients. From this point of view the Web may be of great help, both as a platform to perform controlled online social experiments, and as a repository of empirical data on large-scale phenomena. It is only in this way that a virtuous cycle involving data collection, data analysis, modeling and predictions could be triggered, giving rise to an ever more rigorous and focused approach to language.

It is worth stressing how the contribution physicists, mathematicians and computer scientists could give should not be considered in any way as alternative to more traditional approaches. We rather think that it would be crucial to foster the interactions across the different disciplines cooperating with linguistics, by promoting scientific activities with concrete mutual exchanges among all the interested scientists. This would help both in identifying the problems and sharpening the focus, as well as in devising the most suitable theoretical concepts and tools to approach the research.

References

Baronchelli, A., Felici, M., Loreto, V., Caglioti, E., & Steels, L. (2006). Sharp transition towards shared vocabularies in multi-agent systems. *Journal of Statistical Mechanics*, P06014.

Castellano, C., Fortunato, S., & Loreto, V. (2009). Statistical physics of social dynamics. *Revues of Modern Physics*, *81*, 591–656.

De Beule, J., & Steels, L. (2005). Hierarchy in fluid construction grammar. In U. Furbach (Ed.), *Lecture notes in artificial intelligence: Vol. 3698. KI 2005: advances in artificial intelligence. Proceedings of the 28th German conference on AI*, Koblenz. Berlin: Springer.

De Vylder, B., & Tuyls, K. (2006). How to reach linguistic consensus: a proof of convergence for the naming game. *Journal of Theoretical Biology*, *242*(4), 818–831.

Loetzsch, M., Spranger, M., & Steels, L. (2008a, submitted). The grounded naming game. *Autonomous Robots*.

Loetzsch, M., van Trijp, R., & Steels, L. (2008b). Typological and computational investigations of spatial perspective. In I. Wachsmuth & G. Knoblich (Eds.), *Lecture notes in computer science: Vol. 4930. Modeling communication with robots and virtual humans* (pp. 125–142). Berlin: Springer.

Loetzsch, M., Wellens, P., De Beule, J., Bleys, J., & van Trijp, R. (2008c). *The babel2 manual* (Technical Report AI-Memo 01-08). AI-Lab VUB, Brussels.

Loreto, V., & Steels, L. (2007). Social dynamics: the emergence of language. *Nature Physics*, *3*, 758–760.

Steels, L. (1996). Perceptually grounded meaning creation. In M. Tokoro (Ed.), *Proceedings of the second international conference on multi-agent systems* (pp. 338–344). Menlo Park: AAAI Press.

Steels, L. (2004). Constructivist development of grounded construction grammars. In D. Scott, W. Daelemans, M. Walker (Eds.), *Proceedings 42nd annual meeting of the association for computational linguistics* (pp. 9–19), Barcelona.

Steels, L., & Belpaeme, T. (2005). Coordinating perceptually grounded categories through language: a case study for colour. *Behavioral and Brain Sciences*, *28*, 469–529.

Steels, L., & Bleys, J. (2005). Planning what to say: second order semantics for fluid construction grammars. In A. J. Bugarín Diz, J. Santos Reyes (Eds.), *Lecture notes in artificial intelligence. Proceedings CAEPIA 2005, 11th conference of the Spanish association for artificial intelligence*, Santiago de Compostela. Berlin: Springer.

Steels, L., & De Beule, J. (2006). Unify and merge in fluid construction grammar. In P. Vogt, Y. Sugita, E. Tuci, & C. Nehaniv (Eds.), *Lecture notes in artificial intelligence: Vol. 4211. Symbol grounding and beyond.* (pp. 197–223). Berlin: Springer.

Steels, L., Kaplan, F., McIntyre, A., & Van Looveren, J. (2002). Crucial factors in the origins of word-meaning. In A. Wray (Ed.), *The transition to language*. Oxford: Oxford University Press.

Steels, L., & Loetzsch, M. (2008). Perspective alignment in spatial language. In K. R. Coventry, T. Tenbrink, & J. A. Bateman (Eds.), *Spatial language and dialogue*. Oxford: Oxford University Press.

Steels, L., Loetzsch, M., & Spranger, M. (2008, submitted). *Semiotic dynamics solves the symbol grounding problem*.

Steels, L., & Spranger, M. (2008). The robot in the mirror. *Connection Science*, *20*(2–3), 337–358.

van Trijp, R. (2008). The emergence of semantic roles in fluid construction grammar. In A. D. Smith, K. Smith, & R. Ferrer i Cancho (Eds.), *The evolution of language. Proceedings of the 7th international conference (EVOLANG 7)* (pp. 346–353). Singapore: World Scientific.

Wellens, P., Loetzsch, M., & Steels, L. (2008). Flexible word meaning in embodied agents. *Connection Science*, *20*(2), 173–191.