

An Agent-Based Model for the Role of Short-Term Memory Enhancement in the Emergence of Grammatical Agreement

Javier Vera*

Universidad Adolfo Ibáñez

Keywords

Short-term memory, grammatical agreement systems, language games, language evolution

Abstract What is the influence of short-term memory enhancement on the emergence of grammatical agreement systems in multi-agent language games? Agreement systems suppose that at least two words share some features with each other, such as gender, number, or case. Previous work, within the multi-agent language-game framework, has recently proposed models stressing the hypothesis that the emergence of a grammatical agreement system arises from the minimization of semantic ambiguity. On the other hand, neurobiological evidence argues for the hypothesis that language evolution has mainly related to an increasing of short-term memory capacity, which has allowed the online manipulation of words and meanings participating particularly in grammatical agreement systems. Here, the main aim is to propose a multi-agent language game for the emergence of a grammatical agreement system, under measurable long-range relations depending on the short-term memory capacity. Computer simulations, based on a parameter that measures the amount of short-term memory capacity, suggest that agreement marker systems arise in a population of agents equipped at least with a critical short-term memory capacity.

1 Introduction

What is the influence of short-term memory enhancement on the emergence of grammatical agreement systems in multi-agent language games? Previous work has recently proposed models based on an interesting hypothesis, stressing the role of cognitive effort in the appearance of grammatical agreement systems [7, 8, 21, 22]. *Agreement* systems suppose that at least two words (more generally, two linguistic items) share some features with each other, such as gender, number, or case [9, 10]. For example, in the Spanish nominal phrase (NP) “la casa amarilla” (*the yellow house*) the gender and number of the NP’s head “casa” are repeated on the determiner “la” and the adjectival “-a” suffix.

* Facultad de Ingeniería y Ciencias, Universidad Adolfo Ibáñez. jxvera@gmail.com

In particular, [8] tested the hypothesis that the emergence of a grammatical agreement system arises from the minimization of semantic ambiguity. Indeed, if word order does not convey any meaning, the emergence of a grammatical system of markers solves the problem of how a population of agents culturally develops and self-organizes a common set of markers that allows them to understand utterances with an exponential number of possible interpretations. The price of compositional language is thus computational complexity. For example, within an idealized speaker-hearer interaction, the speaker transmits the utterance “blue pen small red table big,” while the hearer does not know which adjectives are about the same noun. Therefore, the hearer faces the problem of parsing all combinations that fit with the current situation. In order to solve this problem, the introduction of agreement markers as *stickers* may help agents to prune incorrect combinations. In “blue-ta pen-be small-ta red-be table-ta big-be,” the markers “-ta” and “-be” indicate all the words referring to the same referent.

Undoubtedly, grammatical agreement systems have strong relationships with short-term memory capacity. From a neurobiological approach, Aboitiz and colleagues (see, for example, [1–4]) suggest the hypothesis that language evolution has mainly imbricated with an increase of short-term memory [5] skills, which has allowed the online manipulation of words and meanings eventually participating in long-range syntactic dependences, such as grammatical agreement systems. Moreover, the processing of long sentences (without hierarchical recursive dependences), in which head-modifier relations have to be maintained online until the sentence stops being processed, requires large short-term memory networks. Thus, this article agrees with [12, 13] in that short-term memory networks (involved here in the development of simpler forms of syntax) are not likely to be rigid modules, and they should rather be understood as highly plastic structures under experience pressures.

In a related approach, [11] proposes a simple framework to quantitatively measure the load of online processing of those long-range syntactic dependences. Memory load is mainly based on the lineal position of words in a sentence. Thus, the minimization of this measure allows one to explain changes in word order and preferred positions between heads and modifiers.

The aim of this article is to propose a simple (even the *simplest*) multi-agent language game for the emergence of grammatical agreement systems, under measurable long-range relations depending on short-term memory capacity. The hypothesis underlying this article is that the emergence of an agreement system needs at least a minimum amount of short-term memory capacity, defined here as the mechanism for the online maintenance of words and meanings and thus for the establishment of long-range dependences between agreement markers. The methodology is essentially based on computer simulations that describe the dynamics for different values of a parameter of short-term memory capacity.

The work proceeds by introducing the main elements of the language game (Section 2). Section 3 then presents the basic protocol and the results of experiments based on the parameter that measures short-term memory capacity. Finally, a brief discussion about the role of short-term memory in language evolution is presented.

2 The Language Game

The language game, closely based on [7, 8], is played by a population $P = \{1, \dots, p\}$ of agents, interacting about a *world* W that consists in a set of individual objects. Two different objects are described by their features or properties—for example, *blue triangle* and *red triangle*. To talk about objects, agents share a predefined *vocabulary* that associates individual objects and properties of these objects with *words*: nouns (“triangle”) and adjectives (“blue” or “red”). Previous work on language dynamics justifies the assumption of predefining a shared vocabulary on a population of agents [20]. Agents play a *description game* [19], in which at each time step two randomly picked agents observe a scene defined by two referents. The aim of the agents is to produce true descriptions of the referents and their properties.

The language of one agent is defined by sets of grammatical agreement markers associated to each adjective. Each marker m belongs to a finite set M . More precisely, the agent $k \in P$ has a list L_k of pairs $(M^k(adj), adj)$,

$$L_k = \left\{ (M^k(adj), adj) \right\}_{adj} \quad (1)$$

where $M^k(adj) \subseteq M$ is the set of markers that denotes the adjective adj . Put differently, the marker $m \in M^k(adj)$ can be understood as a competing *hypothesis* for the adjective adj .

At each discrete time step $t \geq 0$, two different agents are chosen uniformly at random, which play respectively the roles of *speaker* and *hearer*. The two agents communicate with each other in a context where a subset of the world W is located, the *context* C , and share focus on it (by means of pointing or eye-gazing, in a more realistic scenario). The interaction involves three interrelated behaviors. In the first place, the speaker conceptualizes the context and utters a sentence, with or without grammatical agreement markers, depending on its short-term memory capacity. Secondly, the hearer's behavior mainly consists in parsing the sentence. Finally, both agents *align* [20] their marker sets, in order to increase the chance of understanding in future communicative interactions.

2.1 Speaker's Behavior

First of all, the speaker selects a context C to act as the shared focus of the interaction. For the sake of simplicity, in the experiments reported in this article, C contains two objects, each attributed by two adjectives. C is formed, for example, by the objects *table* and *pen*, each respectively attributed by the properties *blue*, *small* and *red*, *big*.

Secondly, the speaker conceptualizes C and retrieves the set of words that covers the objects and their properties. The speaker creates a random ordering of these words and then utters them in a sentence of an English-like language:

“blue pen small red table big” (2)

Words representing objects are *heads* of an NP. Thus, the sentence (2) contains two disordered NPs. Notice that all the sentences of this article are constructed with six words (two nouns and four adjectives).

Thirdly, the speaker parses his own utterance in order to calculate the amount of short-term memory load (for an alternative hypothesis on why speakers would parse their own utterances, see [18]). The short-term memory load (*SML*) is measured by the length of the syntactic dependencies between a head (the noun that represents one object) and its modifiers (the two adjectives) [11]. More precisely, *SML* reads

$$SML = \frac{\left(\sum_{\text{head}} \sum_{\text{modifier}(\text{head})} d(\text{head}, \text{modifier}(\text{head})) \right) - 4}{10} \quad (3)$$

and it is an increasing function in $[0, 1]$. To define the distance d , imagine that the positions of the heads and their modifiers are specified by natural numbers from 1 to L , where L is the length of the utterance. Following this convention, the utterance “blue pen small red table big” has “blue” at position 1, “pen” at position 2, “small” at position 3, “red” at position 4, “table” at position 5, and “big” at position 6; and $L = 6$. In general, the distance between two elements is defined as the absolute value of the difference between their positions. The distance between a head and one of its determiners is therefore the absolute value of the difference between head's and determiner's positions (see Figure 1 for more details).

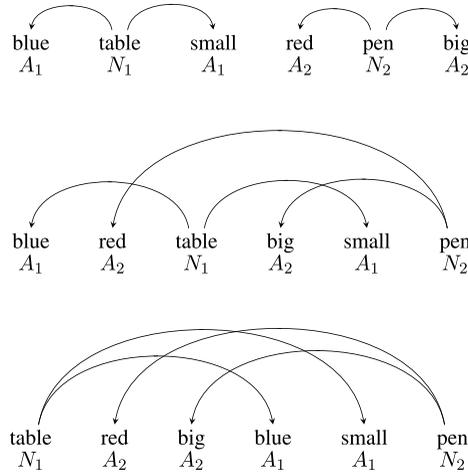


Figure 1. Short-term memory load SME for different dependence trees. Based on the context C , the speaker transmits a sentence constructed by one of the three different linear orderings of the nouns “table” and “pen” and, respectively, the adjectives “blue”, “small” and “red”, “big”. For each ordering, dependence trees are defined by drawing directed edges from each NP’s head to its adjectives. In this case, $SML = \frac{d(\text{table, blue}) + d(\text{table, small}) + d(\text{pen, red}) + d(\text{pen, big}) - 4}{10}$. Top: The lineal ordering “blue table small red pen big” minimizes the short-term memory effort. Indeed, $SML = \frac{1+1+1+1-4}{10} = 0$. Middle: The lineal ordering “blue red table big small pen” is associated to an intermediate short-term memory effort: $SML = \frac{2+2+4+2-4}{10} = 0.6$. Bottom: Finally, the lineal ordering “table red big blue small pen” maximizes $SML = \frac{3+4+4+3-4}{10} = 1$.

Fourthly, the speaker decides, based on the short-term memory load involved in its own sentence (SML), whether to introduce grammatical agreement markers:

- If $SML > \theta$, the speaker introduces (pairwise) grammatical agreement markers, where $\theta \in [0, 1]$ is the global parameter of the agent’s *short-term memory capacity* (the less the parameter θ , the more the short-term memory skills). For example,

$$\text{“blue - ta pen - be small - ta red - be table - ta big - be”}. \tag{4}$$

Since this article considers one-to-one mapping of features, the marker “-ta” can denote either “blue” or “small”, and “-be” can denote either “big” or “red”. The markers “-ta” and “-be”, each one denoting different adjectives, belong to different marker sets. For example, “-ta” $\in M^S(\text{“blue”})$ and “-be” $\in M^S(\text{“red”})$. The markers referring to the adjectives “blue” and “red” must be different. In order to satisfy this requirement, the speaker eventually chooses at random some marker from M . The speaker transmits the sentence (4) to the hearer.

- Otherwise ($SML \leq \theta$), the speaker does not introduce grammatical agreement markers. Therefore, it transmits the sentence (2) to the hearer.

2.2 Hearer’s Behavior and Alignment Strategies

First, the hearer parses the utterance and then diagnoses the short-term memory load. *If the utterance contains grammatical agreement markers*, the hearer identifies each NP’s head and its associated adjectives, by simple inspection of the repeated marker strings. For simplicity, the hearer receives the two marker mappings made by the speaker: The markers “-ta” and “-be” transmit respectively the adjectives “blue” and “red”.

Next, the hearer diagnoses the received marker-adjective associations on its language L_H . Both agents align [20] their languages in order to increase the chance of future successful interactions. For each marker of the transmitted sentence, the hearer follows rules inspired by the *naming game* [6]:

- marker “-ta”
if “-ta” belongs to $M^H(\text{“blue”})$, both speaker and hearer collapse their marker sets associated to the adjective “blue”, that is, they establish, respectively, ($\{\text{“-ta”}\}^S, \text{“blue”}$) and ($\{\text{“-ta”}\}^H, \text{“blue”}$) as their new pairs formed by a marker set and the adjective “blue”.

otherwise, the hearer adds the marker “-ta” to its marker set $M^H(\text{“blue”})$; then it establishes ($M^H(\text{“blue”}) \cup \{\text{“-ta”}\}, \text{“blue”}$) as its new pair formed by a marker set and the adjective “blue”.
- marker “-be”
if “-ta” belongs to $M^H(\text{“red”})$, both speaker and hearer collapse their marker sets associated to the adjective “blue”, that is, they establish, respectively, ($\{\text{“-be”}\}^S, \text{“red”}$) and ($\{\text{“-be”}\}^H, \text{“red”}$) as their new pairs formed by a marker set and the adjective “red”.

otherwise, the hearer adds the marker “-ta” to its marker set $M^H(\text{“red”})$; then it establishes ($M^H(\text{“red”}) \cup \{\text{“-be”}\}, \text{“red”}$) as its new pair formed by a marker set and the adjective “red”.

On the other hand, *if the utterance does not contain grammatical agreement markers*, the hearer does not consider alignment strategies. Nevertheless, a more complex model should suppose short-term memory restrictions for the hearer. For example, the hearer would parse utterances that minimize the short-term memory load ($SML = 0$). Based on the example context C , the hearer would be able to parse the utterances: “small table blue red big pen” or “red big pen small table blue”.

Moreover, hearers would face the problem of what constituent orders allow them to parse sentences with a short-term memory load exceeding a predefined threshold. More generally, each agent would be endowed with a more complex form of syntax, consisting in the association between constructions or *n-grams* [17, 22], differing only in word order, and a system of morphological agreement markers. Alignment strategies would consist therefore in two competing pressures that minimize short-term memory load: (i) the preference for specific word orders [14], and (ii) the development of a system of grammatical agreement markers.

3 Simulations

3.1 Protocol

Three measures describe the dynamics of the proposed language game. In the first place, two related measures are defined: the *average number of markers per agent*,

$$w_T(t) = \frac{1}{p} \sum_{k \in P} \sum_{adj} |M^k(adj)| \quad (5)$$

and the *average number of different markers per agent*,

$$w_D(t) = \frac{1}{p} \sum_{adj} \left| \bigcup_{k \in P} M^k(adj) \right|. \quad (6)$$

The third measure attempts to describe the *amount of alignment of the entire population*:

$$A(t) = \frac{1}{\#adj's} \sum_{adj} \left| \bigcap_{k \in P} M^k(adj) \right|. \quad (7)$$

The analysis is focused on a population P of $p = 10^3$ agents, each one located on a vertex of a complete graph. Agents communicate with each other about 10 nouns, 20 adjectives, and a set of grammatical agreement markers M of size 20. The three measures are bounded: $w_T(t) \leq 20^2$, $w_D(t) \leq 20$, and $0 \leq A(t) \leq 1$. To the extent that $A(t)$ increases, the amount of alignment of the entire population increases. Measures are registered at time steps $p \times t$, with $t \in \{0, 1, 2, \dots, 2p\}$, for several values of the parameter θ , which varies from 0 to 1 with an increment of 5%. The calculations average over 20 initial conditions where each agent $k \in P$ receives a list $L_k = \{(\{m\}^k, adj)\}_{adj}$ and where, for each adj , m is chosen at random from M .

3.2 Results

There are several remarkable results, as shown in Figure 2. First of all, the evolution of $w_T(t)$ and $w_D(t)$ over time exhibits three phases: (i) a fast increase of the number of markers (due to the random selection of markers); (ii) the appearance of a maximum at which $w_T(t) \approx 62$ and $w_D(t) \approx 19$; and (iii) a final decreasing phase of convergence. Phase (iii) evidences the influence of the short-term memory capacity parameter θ . Indeed, to the extent that $\theta > 0.4$ increases, the final values of both $w_T(t)$ and $w_D(t)$ tend to increase. On the other hand, the evolution of $A(t)$ exhibits an S shape: (i) a first stationary value, which depends on θ , followed by (ii) a fast increase that coincides with the decrease of $w_T(t)$ and $w_D(t)$, and (iii) a final plateau, which appears for $t/p > 1500$.

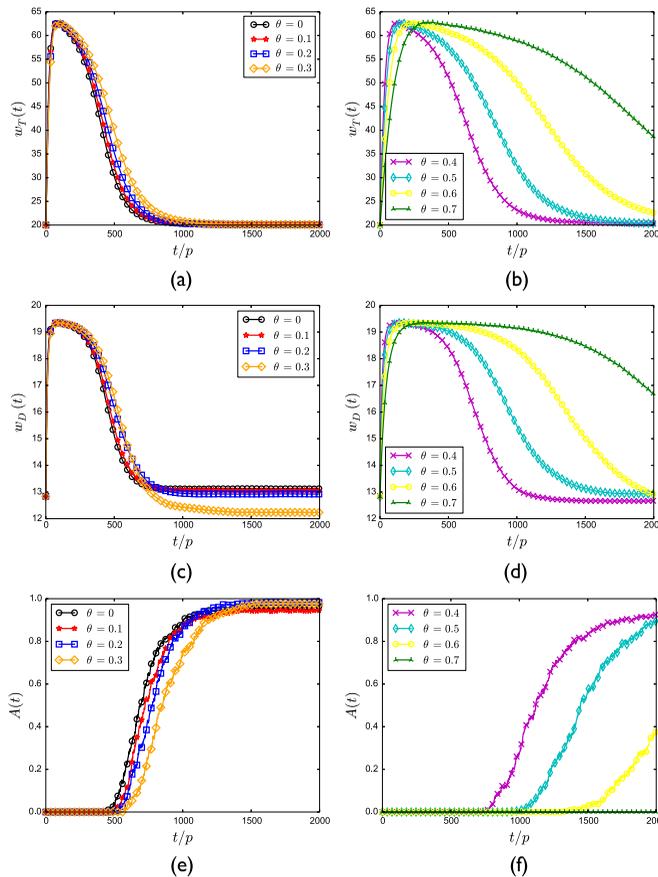


Figure 2. $w_T(t)$, $w_D(t)$ and $A(t)$ versus t , for different values of θ . The simulations run $2p^2$ time steps. Top: $w_T(t)$ versus t ; middle: $w_D(t)$ versus t ; bottom: $A(t)$ versus t . One step t means p speaker-hearer interactions. The parameter θ varies from 0 to 0.7 with an increment of 10%. The calculations average over 20 initial conditions.

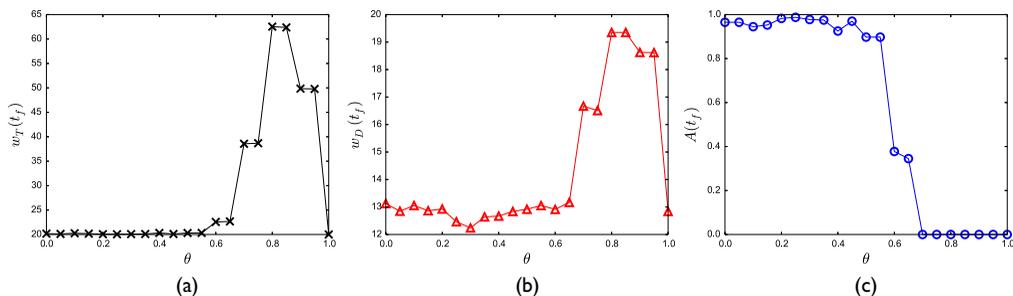


Figure 3. $w_T(t_f)$, $w_D(t_f)$, and $A(t_f)$ for different values of θ . The value of each measure for several values of θ is exhibited after $t_f = 2p^2$ time steps. (a) $w_T(t_f)$; (b) $w_D(t_f)$; (c) $A(t_f)$. The parameter θ varies from 0 to 1 with an increment of 5%. The calculations average over 20 initial conditions.

The observation of the measures after $t_f = 2p^2$ time steps provides a more profound picture of the influence of the parameter θ , as shown in Figure 3. At the critical value $\theta^* \approx 0.6$, the dynamics of the formation of an agreement system exhibits a drastic change. In the first place, for $\theta < \theta^*$ the population culturally develops a shared agreement system (with $A(t_f) \approx 1$) in which each adjective has only one associated marker ($w_T(t_f) \approx 20$) and there are $w_D(t_f) \approx 12$ different markers. Secondly, at $\theta = \theta^*$, the number of markers increases until a maximum, $w_T(t_f) \approx 62$ and $w_D(t_f) \approx 19$, whereas $A(t_f)$ drastically decreases. Finally, for $\theta > \theta^*$, $w_T(t_f)$ and $w_D(t_f)$ decrease and $A(t_f)$ reaches a stationary value close to 0.

4 Conclusion

This brief article discusses an earlier exploration of how and why a grammatical agreement system emerges and culturally propagates, which is based on the hypothesis that in a population of agents those systems arise from the minimization of semantic ambiguity. Nevertheless, the proposal developed here focuses on the fact that sentence processing, particularly the online maintenance of linguistic items while the sentence stops being processed, is closely related to the plasticity (or the enhancement) of the short-term memory networks as a consequence of experience (for a related account, see [23]). This suggests a simple new agent-based approach to study the relationship between short-term memory enhancement and the evolution of grammatical agreement systems. Interestingly, this work demonstrated how in a population of agents the cultural development of simpler forms of syntax (systems of agreement markers) is possible only for agents equipped with at least a critical amount of short-term memory capacity ($\theta \leq \theta^* \approx 0.6$).

Many extensions of the model proposed here should be studied in order to deeply understand the role of cognitive enhancement in the emergence of grammatical devices. In the first place, extensions of the agent-based model could involve possible scenarios of language origins: What happens if two populations of agents with radically different short-term memory capacities interact each other, negotiating grammar? A plausible hypothesis is that agents endowed with more complex memory skills have a greater fitness than other groups with incipient cognitive mechanisms. More formally, on the geographical boundaries of the two populations A and B it is possible, for example, to find a communicative interaction in which the speaker and the hearer belong respectively to A and B , and they are endowed with short-term memory parameters $\theta_A > \theta_B$. Secondly, a study of the relationship between the emergence of recursive structures and short-term memory capacity should be carried out. An interesting startpoint for this task is [15]. Generally speaking, recursion iterates phrases composed of nested phrases, generating hierarchical levels and permitting movement operations [16]. In this connection, recursion supposes the emergence of more complex forms of language, communicating therefore more complex messages; crucially, this allows better cooperation between agents

and thus the establishment of better social bonds [3]. As a preliminary attempt to model the influences of short-term memory enhancement on the development of recursion, one may introduce two parameters controlling short-term memory skills: (i) a parameter α for the maximum number of online items while the sentence is being processed, and (ii) a parameter β for the maximum depth of the hierarchical recursive relationships between linguistic items.

Acknowledgments

The author thanks the anonymous referees for useful commentaries.

References

1. Aboitiz, F. (2012). Gestures, vocalizations, and memory in language origins. *Frontiers in Evolutionary Neuroscience*, 4, 1–15.
2. Aboitiz, F., Aboitiz, S., & García, R. R. (2010). The phonological loop: A key innovation in human evolution. *Current Anthropology*, 51(S1), 55–65.
3. Aboitiz, F., García, R. R., Bosman, C., & Brunetti, E. (2006). Cortical memory mechanisms and language origins. *Brain and Language*, 98(1), 40–56.
4. Aboitiz, F., & García, V. R. (1997). The evolutionary origin of the language areas in the human brain. A neuroanatomical perspective. *Brain Research Reviews*, 25(3), 381–396.
5. Baddeley, A. (2007). *Working memory, thought, and action*. Oxford, UK: Oxford University Press.
6. Baronchelli, A., Felici, M., Loreto, V., Caglioti, E., & Steels, L. (2006). Sharp transition towards shared vocabularies in multi-agent systems. *Journal of Statistical Mechanics: Theory and Experiment*, 2006(06), P06014.
7. Beuls, K., & Höfer, S. (2011). Simulating the emergence of grammatical agreement in multi-agent language games. In T. Walsh (Ed.), *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence, Vol. 1, IJCAI'11* (pp. 61–66). Menlo Park, CA: AAAI Press.
8. Beuls, K., & Steels, L. (2013). Agent-based models of strategies for the emergence and evolution of grammatical agreement. *PLoS ONE*, 8(3), e58960.
9. Boeckx, C. (2008). *Aspects of the syntax of agreement*. New York: Taylor & Francis.
10. Corbett, G. (2006). *Agreement*. Cambridge, UK: Cambridge University Press.
11. Ferrer i Cancho, R. (2015). The placement of the head that minimizes online memory. *Language Dynamics and Change*, 5(1), 114–137.
12. Fuster, J. M. (1999). *Memory in the cerebral cortex: An empirical approach to neural networks in the human and nonhuman primate*. Cambridge, MA: MIT Press.
13. Fuster, J. M. (2009). Cortex and memory: Emergence of a new paradigm. *Journal of Cognitive Neuroscience*, 21(11), 2047–2072.
14. Futrell, R., Mahowald, K., & Gibson, E. (2015). Large-scale evidence of dependency length minimization in 37 languages. *Proceedings of the National Academy of Sciences of the U.S.A.*, 112(33), 10336–10341.
15. Garcia-Casademont, E. (2017). A case study in the emergence of recursive phrase structure. In P. Bourguine, P. Collet, & P. Parrend (Eds.), *First Complex Systems Digital Campus World E-Conference 2015* (pp. 333–336). Cham: Springer International Publishing.
16. Hauser, M., Chomsky, N., & Fitch, W. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298(5598), 1569.
17. Manning, C. D., & Schütze, H. (1999). *Foundations of statistical natural language processing*. Cambridge, MA: MIT Press.
18. Steels, L. (2003). Language re-entrance and the ‘inner voice.’ *Journal of Consciousness Studies*, 10(4–5), 173–185.
19. Steels, L. (2004). Constructivist development of grounded construction grammars. In *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics, ACL '04* (pp. 9–16). Stroudsburg, PA: Association for Computational Linguistics.
20. Steels, L. (2011). Modeling the cultural evolution of language. *Physics of Life Reviews*, 8(4), 339–356.

21. Steels, L. (2016). Agent-based models for the emergence and evolution of grammar. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1701), 20150447.
22. Steels, L., & Casademont, E. G. (2015). Ambiguity and the origins of syntax. *The Linguistic Review*, 32(1), 37–60.
23. Wellens, P. (2012). *Adaptive strategies in the emergence of lexical systems*. Unpublished doctoral dissertation, Vrije Universiteit Brussel, Brussels.