Selective advantages of syntactic language — a model study

Willem H. Zuidema*
jelle@csl.sony.fr
Theoretical Biology, Utrecht University
Padualaan 8, 3584 CH Utrecht
The Netherlands

Paulien Hogeweg
p.hogeweg@bio.uu.nl
Theoretical Biology, Utrecht University
Padualaan 8, 3584 CH Utrecht
The Netherlands

Abstract

We study a computational model of the evolution of language in groups of agents to evaluate under which circumstances syntax emerges. The fitness in the model depends on the composition of the population. We find that this fact significantly alters the evolutionary dynamics. If scores are attributed to both speaker and hearer, expressive syntax is hard to obtain. If scores are attributed only to the hearer, syntax develops, but agents loose the willingness to speak. Implications and a possible solution of this paradox are discussed.

Introduction

Among the many differences between human language and other animal communication systems, syntax is widely acknowledged to be particularly important. Syntax allows us to combine a finite set of meaningful units into an unbounded set of combinations. It allows us to speak about events happening at other times and places. It allows us to communicate about causal relations, to phrase questions or imperatives, and to share in detail previous experiences. The emergence of syntactic language is therefore considered to be one of the major transitions in evolution (Szathmáry & Maynard-Smith, 1995).

In the traditional view, syntax reconciles the need for high expressiveness with some of the natural boundary conditions on communication such as memory limitations, errors in distinguishing sounds, or bottlenecks in the transmission of language knowledge. However, present-day language fulfills many more functions than exchanging information, including facilitating social relations, individual expression, increase of status, esthetic experience and perhaps internalizing our knowledge of the world. It is unclear in what way such functions are recent side-effects, or play an important role in explaining the origins of language.

Discussions of such issues tend to be very unsatisfactory, because they seem hardly restricted by empirical or theoretical bounds. *Computational modeling* offers a novel approach to these issues, because such models are at least restricted by whether or not the *combination* of assumptions implemented in the model yield the hypothesized outcome: syntactic language. This paper discusses a simple computational model of an evolving group of communicating individuals and studies under which selection pressures expressive, syntactic language arises. Before describing the model architecture and results, we will first briefly discuss the theoretical background and some related work.

Evolution of language

Probably the most well-known speculation on the origins of human language is the paper of Pinker & Bloom (1990). Pinker & Bloom argue that syntax must originate in a process of evolutionary optimization, because "natural selection" is the only explanation for the origins of complex design in nature. The paper brings together a valuable collection of findings, but from a theoretical perspective it is problematic, because it lacks precision and formalization. In its weakest interpretation the central claim is trivial (there is no doubt that only members of the human species can acquire fluency in a human language) and in its strongest interpretation ("evolution has led to genes that explicitly specify a universal rule system for language") the claim is untenable. However, the lack of a more precise aspect to Pinker & Bloom's work, makes it hard to position their ideas between these extremes.

Moreover, Pinker & Bloom's paper is symptomatic for the popular fallacy in linguistics that one can only choose between two explanations: (i) language originates in a genetic evolution, or (ii) language arises as the spontaneous result of general cognitive skills and social structure. We believe that putting these two explanations in opposition, excludes the most interesting part of the story. Spontaneous pattern formation ("self-structuring") needs a mechanism to set the right parameters, and evolution needs a plausible substrate to operate on. Viewing self-structuring as a substrate for evolution (Boerlijst & Hogeweg, 1991a) offers a fresh perspective that allows one to study how evolution, genetic information, learning, development, embodiment and social structures all interact to shape

*Present address: Sony CSL, 6, Rue Amyot, 75005, Paris, France; webpage: www-binf.bio.uu.nl/~jelle human language. Note that such an interactionist account differs fundamentally from a naive "some parts of language are innate, some are learned" view.

Computational modeling

Recent work that studied such interactions in computational models has produced a wealth of new hypotheses and insights (Hurford, 1989; Hashimoto & Ikegami, 1996; Batali, 1997; Steels, 1997; De Boer, 1999; Kirby, 2000; Nowak & Krakauer, 1999; Hurford, 2000). Such models are *relatively precise* implementations of the underlying set of assumptions, and allow one to evaluate the internal coherence of such a set. Moreover, they are *productive*, in the sense that they often show unexpected behaviors that help to generate new hypotheses and concepts. And although they are necessarily simplified representations, the fact that their behavior can be experimentally evaluated makes it possible to study more complex phenomena than with analytical methods alone. Computational models therefore pre-eminently can make tractable systems with many variables and interactions.

On the issue of the origins of syntax, a number of intriguing mechanisms have been identified using computational modeling techniques. Although very diverse, they all emphasize the fact that syntax greatly increases the number of possible forms in a language. For instance, Batali (1997), Kirby (2000) and Hurford (2000) studied how *cultural evolution* can account for the emergence of syntax. Although they use several different formalisms, the common idea in this work is that the internal knowledge of language (the infinite "I-language") is transmitted culturally (via a finite "E-language") from one agent to another. This "transmission bottleneck" works as a filter, in which syntactic elements of language typically out-compete non-syntactic elements, because the former are inherently used more often.

Nowak & Krakauer (1999) studied a game-theoretic model of language evolution and identify a different mechanism that can account for the emergence of syntax. Using the matrix representations of Hurford (1989), they infer a "linguistic error limit". Given that an individual makes mistakes in distinguishing sounds with a probability that depends on the similarity between those sounds, Nowak & Krakauer calculate a limit on the number of messages an individual can convey. They show mathematically that *word formation* and *syntax* can help overcome such a limit. Moreover, they show that both non-syntactic and syntactic strategies are *evolutionary stable strategies* (i.e. cannot be invaded by other strategies). However, every mixed strategy can be invaded by every mixed strategy that uses *more* syntactic sentences. Thus, the evolutionary process should lead towards grammar.

Hashimoto & Ikegami (1996) showed that syntax can emerge in an *evolving group* of communicating agents. The agents in their model have an internal rewriting grammar, that generates a formal language using lexical or syntactic strategies. Because there is no limit on the number of rules, both strategies could in principal generate all possible strings in the finite domain that was used. However, at the start of the simulations agents are initialized with just one rule in their grammar. Because mutations add rules one at a time, and expressiveness grows much faster with grammar size using a syntactic strategy, syntactic agents outcompete non-syntactic ones.

An important aspect of Hashimoto & Ikegami's model is that fitness is not a fixed measure, but depends on the kind of grammars that are present in the population. This leads to some counterintuitive results. For instance, they find that the most expressive agents are not necessarily the most successful and that a score for *not being recognized* accelerates the evolution of syntax. These observations are the starting point for the model study reported in this paper.

The model

The model reported in this paper is a variant of the model of Hashimoto & Ikegami (1996). Of the many aspects that might be relevant, we study only one particular type of interaction: between evolutionary dynamics and group structure. We therefore ignore all aspects of grammar, except for the fundamental properties of compositionality and recursion. We ignore semantics, by just attributing scores for successful parsing. And we ignore learning, by assuming that agents end up with the same internal grammar, except from some changes that result from mutations in the innate component of language.

In this simplified model we will show that evolution shapes the linguistic environment of agents, but, conversely, that the group structure also shapes the evolutionary process. This interaction guides evolution in unexpected directions, and, depending on the implemented function of language, can both facilitate and hinder the development of syntax.

The model consists of a population of agents with an internal rewriting grammar, which they inherit with some mutations from their parent. The grammars are context free grammars, with nonterminal and terminal symbols from the small alphabets $V_{nt} = \{S, A, B\}$ and $V_{te} = \{0, 1\}$ respectively. As an extra restriction, we don't allow the "S" at the right-hand sides of rules. At the start of most simulations, agents are initialized with a grammar with just one rule: randomly $S \mapsto 0$ or $S \mapsto 1$. Agents have the ability to derive ("speak") and parse ("understand") strings of 0's and 1's of maximum length 6, using the rules from the grammar. Within these constraints the maximum expressiveness is 126. We define compositionality as using the non-terminals A and B, and recursion as using rules that were used before in the same branch of the rewriting tree.

Agents interact in a set-up of "language games". In every game all agents can speak one string and try to recognize the strings produced by other agents. Every generation a number of games is played and scores are attributed for successful communication. In most simulations, we use an explicit "innovation pressure". This pressure is implemented by discounting scores with the number of times a string is already heard before, and corresponds to a semantic need for a rich repertoire of forms. We designed several scoring schemes that reflect hypotheses on the function of language. The most important schemes are labeled "communication" and "perception":

communication corresponds to a selection pressure to optimize the total of exchanged information, such that both the speaker and the hearer benefit from successful communication. This pressure is implemented by a score for recognition and for being recognized.;

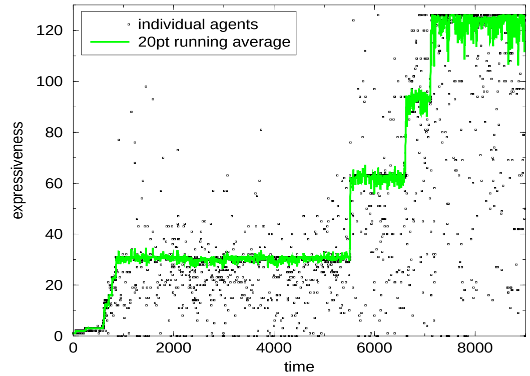
perception corresponds to a selection pressure to optimize the total of information received, in order to make use of the knowledge of others (as if one indirectly shares someone else's *perception*). This pressure is implemented as a score for recognition;

We replace all agents every generation with offspring of the present population. The number of offspring of an agent depends on the total score it has received relative to other agents. Random mutations are applied to the offspring with fixed probabilities for modification of existing rules ("replace"), duplication of a random rule ("add") or deletion of a rule ("delete"). We also implemented a mutation "shift", that swaps a rule with the previous rule in the grammar and occurs with a probability per rule. These mutations correspond to conventions in evolutionary programming and allow for optimizing some of the relevant features of grammars, but otherwise they are more or less arbitrary.

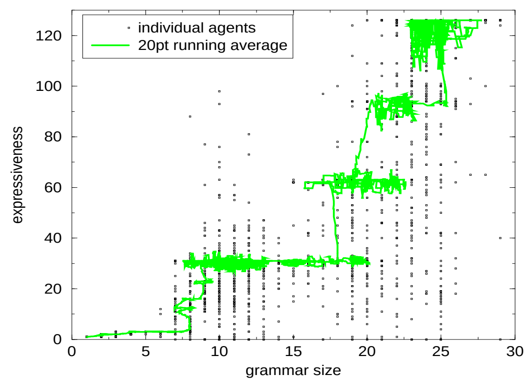
The group effect

Fitness in this model is not a static function of an agent's grammar ("genotype"), but it depends on the grammars of other agents too. The general observation in experiments with the model with many different parameter settings is that this fact strongly influences the evolutionary dynamics (Hashimoto & Ikegami, 1996). The success of an agent's individual language is determined by how well it matches the language of the whole group, rather than by how much information it can encode ("expressiveness"). We call this phenomenon the "group effect".

Figure 1(a) shows an example simulation, with a "communication" scoring scheme and "innovation", that shows clearly some of the mechanisms that play a role. From the initial level of expressiveness of 1, the



(a) Expressiveness over 9000 generations



(b) The same run in a "phase space"

Figure 1: *An example run with very clear epochal evolution. Shown are the running averages and individual agents at every tenth generation. Note that most individual points are hidden under the grey line. (a) During an epoch, expressiveness stays at a fixed level. In fact, in the first stage (E=31) the dominant language stays exactly the same for thousands of generations. Individual agents with higher expressiveness occur, but are not able to survive in the group. (b) Grammars do vary, however, which is possible because of the neutrality in the grammar–language mapping (see text). In the phase space, one can clearly see that grammar size fluctuates during an epoch. All jumps to higher levels take place when grammars are relatively large. Such grammars are clearly larger than necessary and have a neutral tails. Parameters: default "communication" run with innovation pressure (see section "selective advantages")*

population evolves within several hundreds of generation to a level of 31. At this point, evolution has developed via selection and random mutations grammars that are redundant and not very structured, and combine several strategies in the rewriting process from the start symbols "S" to a distinct sequence of terminal characters.

For a very long time, from around generation 860 until 5510, the population remains fixed at a level of expressiveness of 31. Analysis of the language reveals that the set of strings of the majority of agents remains unchanged for this whole period. However, frequently agents appear that have a much higher level of expressiveness. This illustrates that (i) the mapping from grammar to language is very non-linear, because a single mutation can make a dramatic change in the size of the language, and (ii) there is a very strong group effect, because agents that have a much higher expressiveness (and thus are "objectively" much better), can nevertheless not persist in the population. The reason is that the languages of these agents differ too much from the language of the group. The agents therefore obtain fewer scores for being recognized and possibly even for recognizing.

Another striking observation in this simulation is that, although the languages remain unchanged for several thousands of generations, the grammars undergo a constant reorganization. This illustrates that the mapping from language to grammar is not only non-linear, but also very redundant.

Figure 1(b) shows a graph of the same simulation in a "phase space" that shows the average grammar size versus the average expressiveness at each generation. As one can clearly see, once a certain level of expressiveness is reached, the evolutionary process "wanders around" for a long time, without significant changes in the expressiveness ("neutrality"). Only when the grammars are relatively large, and thus have many unused, redundant rules, a chance event causes the population to jump to a new level of expressiveness. This chance event is that two agents mutate to the same richer language, and thus can obtain in their mutual communication enough scores to compensate for differing from the group. This mechanism relates to the idea of "neutral networks" — networks of connected points in genotype space that correspond to the same phenotype — that forms a good explanation for the occurrence of "epochs" or "punctuated equilibria" in evolving systems with a fixed fitness function (Van Nimwegen *et al.*, 1999).

Selective advantages

While the "group effect" occurs under all parameter settings of the model, its role can be quite different for each of the scoring schemes and the initial grammars we considered. We observe compositional and recursive grammars only in about half of the parameter combinations we considered. Even if scores are explicitly discounted with the number of times a string is already used before ("innovation pressure"), expressive syntax does not necessarily emerge.

This fact is surprising, because the intuitive expectation is that expressiveness is selectively advantageous. Indeed, with (i) an *explicit* innovation pressure, the

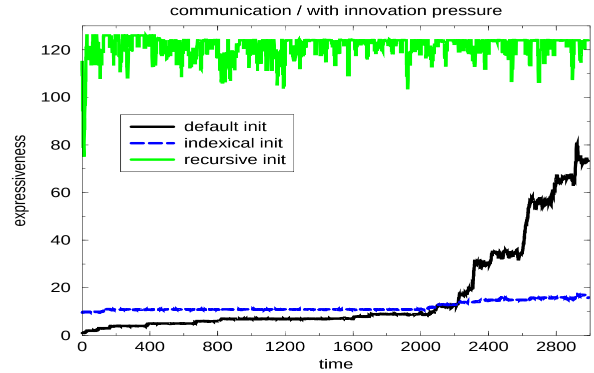


Figure 2: *Communication with innovation pressure for three different types of initial grammars. With a sufficiently large initial lexical grammar, expressive syntax can not develop.*

average score per agent has it optimum at maximal expressiveness. However, *implicitly* expressiveness influences the scores in other ways as well: (ii) expressive speakers are more likely not to be understood, and (iii) expressive listeners are more likely to understand.

This leads to an interesting interplay between each of these roles of expressiveness and the group effect. Under communication settings (ii) not being recognized is *disadvantageous*, while (iii) recognition is *advantageous* and in both scoring dimensions similarity to the group's language is important. Under perception settings (ii) not being recognized and (iii) recognition are beneficial, while similarity to the group's language is important for recognition, but *dissimilarity* is better for not being recognized (and thus hindering one's competitors). Moreover, the strength of the group effect depends on the size of the group's language and the variation within the group. In various experiments we obtained the following results:

communication does not lead to highly expressive grammars with the default initial grammar and without the innovation pressure. If the initial grammar is an expressive, recursive grammar, the high level of expressiveness can be maintained. In contrast, with a medium size lexical grammar, grammars remain lexical and expressiveness remains limited.

With an innovation pressure and the default initialization expressive syntax eventually does develop. In this type of runs we observe a stepwise development, with typically long intervals at the same level of expressiveness. Expressive syntactic grammars are reached only after very many generations. With an expressive, recursive initial grammar, the high level of expressiveness can be maintained. With a medium size lexical grammar expressiveness remains limited and no syntax develops (see figure 2).

With "communication" as the function of language,

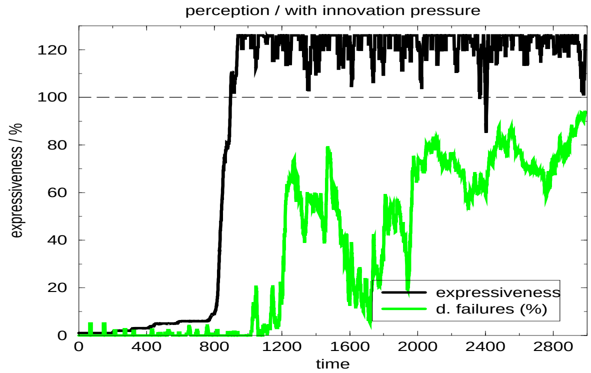


Figure 3: *A typical example of a simulation with "perception" settings, the default initial grammar and an innovation pressure. Shown are the average expressiveness over time, and the percentage of failures in derivation. After around 3000 generations this percentage approaches 100, indicating that very little communication is maintained.*

syntax can thus be maintained if present, but is hard to obtain. If the initial grammar is of sufficient size and of a lexical type, syntax never develops. These results are particularly interesting, as they resemble the situation that is traditionally thought to precede the emergence of grammar: large, lexical protolanguages, with communication benefits for both speaker and hearer.

perception shows rapid growth in expressiveness in most cases considered. With the default initialization and no innovation pressure, expressive syntax develops within a few hundred generations. With the lexical initialization it takes much longer, but the development of syntax was usually observed.

With an innovation pressure and default initial grammars the growth is generally slower than without such an innovation pressure. Infrequently, we even observe runs that remain lexical throughout the simulation. When initialized with an lexical grammar, the runs with innovation pressure show such behavior.

"Perception" thus yields expressive syntax in most cases considered (see figure 3). The benefits of *not being understood* seem to be a strong incentive to develop more expressive language. Interestingly, an innovation pressure makes the development of syntax *less* likely. Apparently, the fact that the hearer benefits from richer input hinders this development.

Paradox

Another striking feature of perception runs is the high number of failures that occur in derivation (see figure 3). Apparently, agents develop grammars that are able to parse a high number of strings, but nevertheless frequently fail in derivation. This is possible because of the asymmetry in parsing (complete bottom-up search of the derivation tree) and derivation (random top-down walk). This possibility was not implemented intentionally. Nevertheless, the evolutionary process discovered it and "actively" exploits it.

This observation points at a important assumption in the model: agents are forced to participate in the language game. A classic altruism problem thus arises: if speaking behavior is beneficial only for an individual's competitors, why would it be retained in evolution? We extended the model with a parameter for probability to speak. Under perception settings this parameter indeed quickly evolves to zero.

Interestingly, these results constitute a paradox: under those circumstances that syntactic expressiveness develops, willingness to speak disappears. Under the circumstances where willingness to speak is retained, syntactic language does not develop. We studied a possible solution for this paradox in a model where agents are localized on a 2D grid and interact only with their immediate neighbors. Such spatial models are known to naturally yield altruism, because spatial patterns make multilevel evolution possible and kin selection more likely (Boerlijst & Hogeweg, 1991b).

The willingness to speak can be retained in the spatial model with perception settings. The parameter that determines the probability of an agent to speak at its turn in the language game, is initialized at 0.1. As one can see in the example of figure 4, the average value rapidly evolves to a high value close to the maximum. Spatial patterns are responsible for this selection pressure towards altruistic behavior. If one destroys the spatial patterns, also the willingness to speak disappears.

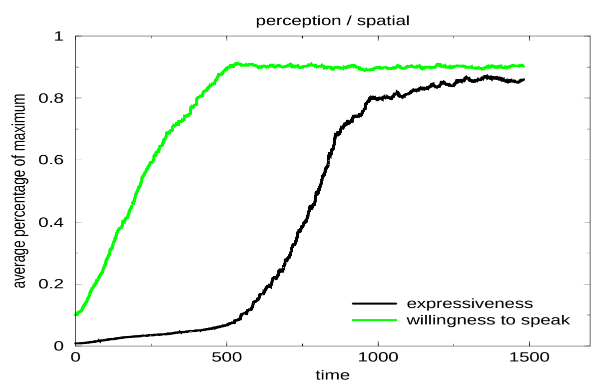


Figure 4: *Perception in space. Shown is the average fraction of the maximum of expressiveness (maximum is 126) and willingness to speak (maximum is 1). Parameters are: initial population size = 2000, number of games per generation = 1, maximum string length = 6, minimum number of understanders = 0, madd = 0.1, mrep = 0.01, mdel = 0.01, maximum number of parsing steps = 500, maximum number of derivation steps = 60, self-interaction not allowed, discount factor 1.0, scores proportional to string length*

Discussion

Some of the striking differences in the results of different scoring schemes can be better understood by looking at a very simple game-theoretic model, where there are just two agents and two levels of expressiveness. If we work out the language games that take place in such a set-up, we find that both the low/low and the high/high situations are equilibria in the communication case, but in the case of perception only the high/high situation is an equilibrium. These results qualitatively corresponds to the results we obtained in the simulations.

The essential observation here is that, although homogeneous high expressiveness is the "best" solution, *unilateral* high expressiveness under communication setting is in fact disadvantageous. It seems a promising approach to extend this game-theoretic analysis to a more general case, with more levels of expressiveness and more interacting agents. However, many aspects of the model behavior depend on the non-linear mapping between grammar and language and can not easily be captured in such an analysis.

Conclusions

Traditionally the origins of language are thought to be explained as either the spontaneous result of human cognitive abilities and social interactions, or the result of an evolution of our innate language capacity. This model study shows an example system where both social interaction and evolutionary updating play a role. Not because one part of language can be explained by "nurture" and another part by "nature", but because they fundamentally interact: social interactions shape the evolutionary process and vice versa.

Also, traditionally language and the evolution of language are studied in terms of how much information about the outside world can be transmitted. Our results suggest that this might not always be the most interesting way of looking at language, because language can have its own dynamics within a group that is quite independent from how well it represents the outside world.

Moreover, this model study shows results that deviate from the traditional picture that lexical protolanguages became larger and larger until syntax became necessary. If communication if beneficial for both speaker and hearer and the population uses an extensive lexical language, syntax does not develop. If the traditional picture holds, the question arises which mechanisms are responsible for the differences.

Finally, spatial patterns have not played much of a role in speculations about the origins of language. Results from this study suggest that such spatial patterns can be relevant. The fact that present-day language shows obvious spatial patterns indicates that a global approximation perhaps excludes important mechanisms.

Many open questions remain. For instance, under perception settings there is an *indirect benefit* of speaking that leads to high values of the willingness to speak. Why then, does this indirect benefit not result in the same disadvantage of unilateral high expressiveness that we observe under communication settings? Such intriguing issues are left for future work.

Acknowledgements

Many thanks to Ludo Pagie for help with programming C++, and Michael Moortgat, Onno Zoeter, and all members of the Theoretical Biology group for technical and moral support and stimulating discussions. WHZ thanks Sony CSL-Paris for allowing him the time to write and present this paper.

References

BATALI, J. (1997). Computational simulations of the emergence of grammar. In: *Approaches to the evolution of language* (Hurford, J. et al., eds.). Cambridge University Press.

BOERLIJST, M. C. & HOGEWEG, P. (1991a). Self-structuring and selection. In: *Artificial Life II* (Langton, C. et al. eds.), 255–276.

BOERLIJST, M. C. & HOGEWEG, P. (1991b). Spiral wave structure in pre-biotic evolution. *Physica D* **48**, 17–28.

DE BOER, B. (1999). *Self-Organisation in Vowel Systems.* Ph.D. thesis, Vrije Universiteit Brussel AI-lab.

HASHIMOTO, T. & IKEGAMI, T. (1996). The emergence of a net-grammar in communicating agents. *BioSystems* **38**, 1–14.

HURFORD, J. (1989). Biological evolution of the saussurean sign as a component of the language acquisition device. *Lingua* **77**, 187–222.

HURFORD, J. R. (2000). Social transmission favours linguistic generalization. In: *The evolutionary emergence of language* (Knight, C. et al., eds.). C.U.P.

KIRBY, S. (2000). Syntax without natural selection. In: *The evolutionary emergence of language* (Knight, C. et al., eds.). C.U.P.

NOWAK, M. A. & KRAKAUER, D. C. (1999). The evolution of language. *Proc. Nat. Acad. Sci. USA* **96**, 8028–8033.

PINKER, S. & BLOOM, P. (1990). Natural language and natural selection. *Behavioral and brain sciences*

STEELS, L. (1997). Synthesising the origins of language and meaning. In: *Approaches to the evolution of language* (Hurford, J. et al., eds.). C.U.P.

SZATHMÁRY, E. & MAYNARD-SMITH, J. (1995). The major evolutionary transitions. *Nature* **374**, 227–232.

VAN NIMWEGEN, E., CRUTCHFIELD, J. & HUYNEN, M. (1999). Neutral evolution of mutational robustness. *Proc. Nat. Acad. Sci. USA* **96**, 9716–9720.