

1 **Group 4: What can mathematical, computational and robotic**
2 **models tell us about the origins of syntax?**

3 Herbert Jaeger, Rapporteur

4 Andrea Baronchelli, Ted Briscoe, Morten H. Christiansen, Tom Griffiths,
5 Gerhard Jäger, Simon Kirby, Natalia L. Komarova, Peter J. Richerson,
6 Luc Steels and Jochen Triesch

7
8
9 **Introduction**

10 Mathematical and computational models play a crucial role in all sciences and can clearly be
11 helpful to the study of language evolution as well. A model makes certain theoretical
12 assumptions about evolutionary forces and linguistic representations and then shows what the
13 consequences of these assumptions are on the outcome of language evolution. Of course, the
14 model does not in itself prove that the assumptions are empirically valid but it shows whether the
15 assumptions are coherent, have the effect they are believed to have, and are in principle
16 sufficient to generate the phenomena they are claimed to generate. Thus, they provide the
17 opportunity to test different hypotheses about the ingredients that are necessary for languages
18 with certain properties to emerge through processes of transmission and interaction between
19 agents. In turn, they make predictions about the relationship between language acquisition,
20 communicative interaction, and language change that can be assessed through experiments with
21 human participants or robotic agents and through comparison with historical data.

22 In the field of language evolution, a wide range of models has already been explored, but this is
23 only the beginning. The complexity of the models, the questions they address, and the techniques
24 used to check the validity of current models vary widely. Consequently, the discussions in our
25 group did not and could not be expected to lead to a unified and complete picture, partly because
26 researchers have been looking at entirely different aspects of the enormously complex problem
27 of language evolution and have been using very different methods. Instead the group tried to
28 sample the landscape of existing modeling efforts (section 1) and the representations of grammar
29 and grammatical processing that are used in them (section 2). We then surveyed arguments
30 regarding why and how modeling can contribute to the overall language evolution research
31 enterprise (section 3), and outline future research including possible collaboration with biologists
32 and linguists (section 4).

33 To avoid a possible misunderstanding, we point out that the discussions in our group, and
34 consequently the materials of this chapter, mostly concern the *cultural* evolution of language, to
35 be distinguished from the *biological* evolution which is in the focus of other contributions in this
36 volume. Nonetheless, investigations of cultural language evolution have implications for
37 research on biological evolution, because if it is found that certain traits of language can
38 naturally be explained by the former, biological mechanisms are relieved from an explanatory
39 load. Conversely, biologically evolved, generic, non-linguistic information processing
40 capabilities (e.g. sequential processing mechanisms) yield the scaffolding for cultural evolution.

1 Another preparatory remark is that modeling efforts adopt approaches that are quite standard in
2 other domains of complex systems science, but may be relatively new to linguists. For example,
3 there is often an effort to seek simplified models in order to clearly pin down the assumptions
4 and, in many cases, make the models tractable from a mathematical point of view. Modelers
5 typically focus on replicating statistical distributions of language phenomena rather than
6 matching directly the particulars of a given human language. They will first consider
7 communication systems that have only a rudimentary resemblance to language before increasing
8 the complexity further step by step. Or they will make assumptions about certain aspects of
9 language interaction (such as joint attention or perception) in order to make simulations doable at
10 all. Some models are not about language per se but address the preconditions for language, such
11 as cooperation (Richerson & Boyd, this volume). It is therefore important to keep in mind that
12 the modeling work discussed here is primarily concerned with investigating the consequences of
13 hypotheses rather than trying to model in detail and in a realistic way the origins and evolution of
14 human language.

15 **1. Paradigms for Studying Language Evolution**

16 The discussions in the group arose from the multi-faceted experience of the participants with
17 computer-based simulations of language dynamics, robotic experiments, and mathematical
18 analysis. We are not aware of any generally accepted way of characterizing or classifying
19 computational modeling approaches in the natural or social sciences. In the present context, we
20 could nevertheless identify a number of different modeling paradigms that have grown up
21 historically based on the shared interests of the researchers involved in exploring them. Each
22 paradigm frames the process of language evolution in a particular way, focuses on some of the
23 forces that might play a role, and then examines specific fundamental questions through concrete
24 models and experiments. Within each paradigm we have seen the development of mathematical
25 models, computational or robotic experiments, and psychological experiments with human
26 subjects. Of course, the distinctions between paradigms that are made here is to some extent
27 arbitrary and not always clear-cut. There are continuous dimensions linking these paradigms and
28 hence considerable opportunities for cross-fertilisation. Moreover, we anticipate that additional
29 modeling paradigms may spring up in the future to explore other aspects of the vast research
30 domain of language evolution.

31 A first distinction that can be made is between agent-based models, which try to pin down the
32 cognitive and social processes that could give rise to forms of language, and macroscopic
33 models, that aggregate the behavior of a population and then formulate equations defining the
34 evolution over time among these aggregate quantities. Another dimension for categorising the
35 models concerns the importance given to cultural transmission, cognition, or biology, which has
36 given rise to Iterated Learning models, Language Games, and genetic evolution models.

37 In Figure 1 we give a schematic illustration of two main dimensions on which the paradigms
38 differ.

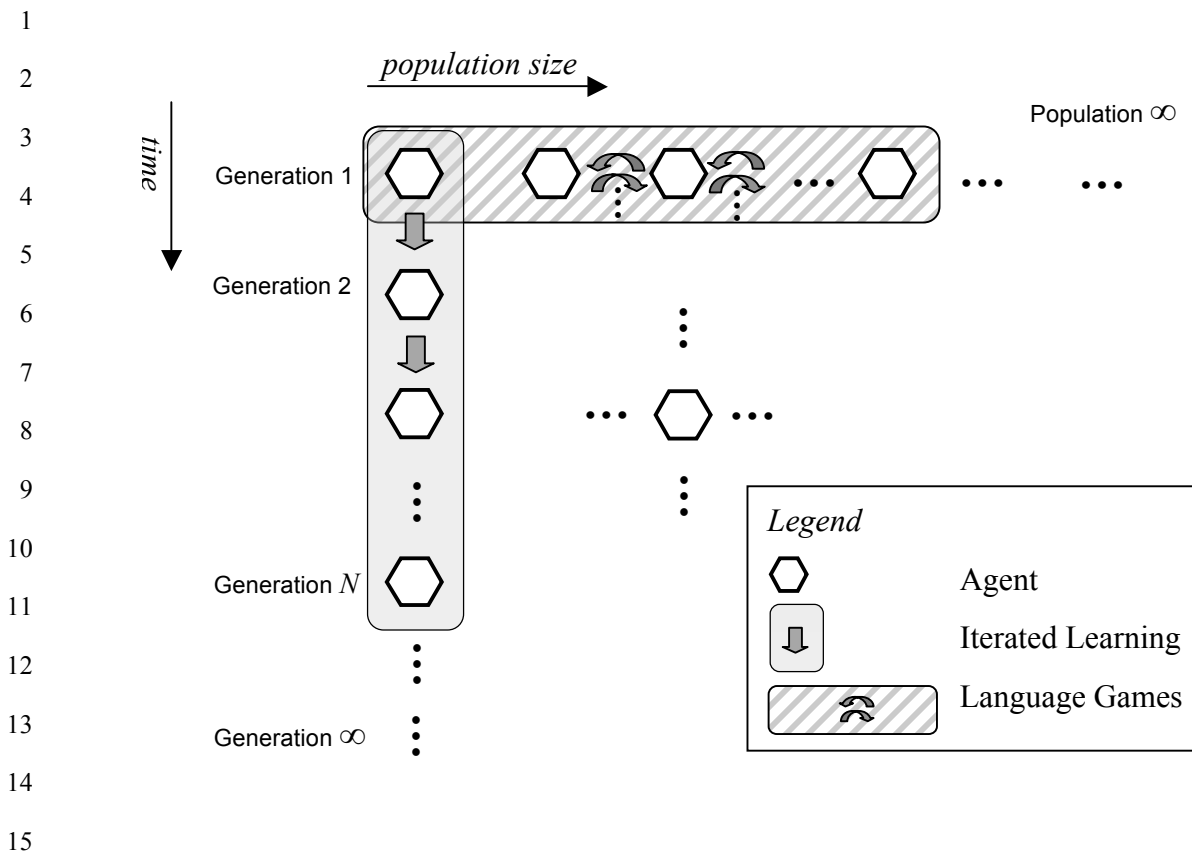


Figure 1. Schematic "coordinate system" comparing for agent-based language evolution paradigms. The simplest models within the Iterated Learning paradigm focus on transmission across *generations* of agents in a singleton chain of teacher-learner dyades; Language Games focus on how language constructs emerge and evolve in interactions between agents. Numerous other paradigms can be seen as mixtures and ramifications of these two.

1.1. Agent-based models

Agent-based models center on models of individual language users as members of populations. The agents are given certain cognitive capabilities (for example a particular learning strategy) and made to interact, for example in the simulation of a teacher-learner situation or a communicative interaction between two individuals. By simulating the effect of a large number of interactions, agent-based models can study under what conditions language systems with similar properties as human natural languages can appear. Agent models vary greatly in complexity, ranging from simple statistical "bag of words" language models to robots using complex grammatical and semantical representation formalisms to communicate with each other in a dynamical environment.

Three types of agent-based models have been developed: iterated learning models which focus on understanding the role of cultural transmission, language game models which emphasise the role of communication and cognition, and genetic models which explore the role of biological evolution.

1 *Iterated Learning*

2 The first paradigm that has already been explored quite deeply is known as the Iterated Learning
3 Paradigm. It focuses on understanding the relationship between properties of the individual and
4 the resulting structure of language by embedding a model of an individual learner in a so-called
5 "transmission chain" (also sometimes called "diffusion chain", see Kirby, Christiansen, &
6 Chater, this volume; Briscoe, this volume, for further details, and Mesoudi, 2007, for a review of
7 this approach to studying cultural evolution more generally). In these models, the linguistic
8 behavior of one individual becomes the learning experience of another individual who in turn
9 goes on to produce behavior that will be input for a third individual and so on. The focus of this
10 framework is on the contribution of learning in shaping the process of cultural transmission, with
11 the goal of specifying precisely the relationship between constraints and biases provided by
12 biology and the universal properties of linguistic structure. The idea is that a fundamental
13 challenge for language is to be repeatedly transmitted between individuals over generations, and
14 the transmission process is imperfect in important ways (e.g., learners have particular biases,
15 they only see a subset of the language, there is noise in the world, and so on). The result is an
16 adaptive system whereby language evolves culturally in such a way to give the appearance of
17 being designed for transmission fidelity.

18 The main simplification in many (but not all) of the models of this "iterated learning" process is
19 that the transmission chain consists of a single individual at each generation, and involves only
20 vertical transmission (i.e., transmission between generations). This simplification allows
21 researchers to focus on the sole contribution of the learning bias plus the nature of the selection
22 of training data (e.g., number of examples, etc.), although it leaves out many of the factors
23 associated with horizontal transmission (e.g., selection of models to learn from, having shared
24 communicative goals, and population structure). One avenue for future research is to explore the
25 implications of other, more realistic models of populations, while maintaining the emphasis on
26 the role of transmission in shaping language structure. For a recent review of general cultural
27 evolution models see McElreath and Henrich (2008).

28 Examples of iterated learning models are given in Kirby, Christiansen and Chater (this volume).
29 An emphasis in many of these models so far has been the explanation of the emergence of
30 compositional structure in language. Compositionality, along with recursion, is the fundamental
31 feature of human syntax that gives us open-ended expressivity. It is also arguably absent in any
32 other species, despite the prevalence of communication in nature. Accordingly, it is an important
33 target for explanation for those interested in the evolution of language. Using mathematical,
34 computational, and experimental models, researchers have examined the conditions under which
35 compositionality and the relationship between compositionality and frequency may emerge.
36 Specifically, these models suggest that compositionality arises when there is a "bottleneck" on
37 the cultural transmission of language - in other words, where learning data is sparse.

38 *Language Games*

39 The second class of models investigates the role of embodiment, communication, cognition and
40 social interaction in the formation of language. Instead of modeling only teacher/learner
41 situations as in iterated learning approach, it models the communicative interactions themselves
42 in the form of language games. A language game is a situated embodied interaction between two

1 individuals within a shared world that involves some form of symbolic communication. For
2 example, the speaker asks for "a cup of coffee" and the hearer gives it to her. When speaker and
3 hearer have shared conventions for solving a particular communicative problem they use their
4 existing inventory in a routine way. But when this is not the case, the speaker requires the
5 necessary cognitive capabilities to extend his inventory, for example expanding the meaning of a
6 word or coercing an existing word into a new grammatical role, and the hearer requires the
7 ability to infer meanings and functions of unknown items and thereby expand his knowledge of
8 the speaker's inventory.

9 In typical language game models, the individuals playing language games are always considered
10 to be members of a population. They interact only in pairs without any centralized control or
11 direct meaning transfer. There is unavoidable variation in the population because of different
12 histories of interaction with the world and others, but a proper selectionist dynamics,
13 implemented by choosing the right alignment and credit assignment strategies for each
14 individual, causes certain variants to be preferred over others. Language game models often
15 operate with a fixed population because they examine the thesis that language emerges and
16 evolves by the invention, adoption, and alignment strategies of individuals in embodied
17 communicative interactions, but many experiments have been done as well in which a flow is
18 organized in the population with members leaving or entering the population, in order to show
19 that the model handles cultural evolution as well.

20 By now there have been dozens of experiments in language games exploring how different
21 aspects of language may arise (see Steels, this volume). The simplest and earliest game studied is
22 the Naming Game, in which agents draw attention to individual objects in the world by using
23 (proper) names (Steels, 1995). Guessing games have been used to study the co-evolution of
24 perceptually grounded categories and words (Steels & Belpaeme, 2005), flexible word meanings
25 (Wellens, et.al., 2008), and the emergence of spatial language (Steels & Loetzsch, 2008).
26 Description games have been used in experiments in the emergence of grammar, particular case
27 grammar (VanTrijp, 2008).

28 Language games have been explored further from three angles: through mathematical analysis,
29 particularly using the methods of statistical physics, through computational simulations and
30 robotic experiments, and through experiments with human subjects as carried out by Galantucci
31 (2005), Pickering & Garrod (2004), and others. Robotic experiments are particularly useful if
32 one wants to study the question how embodiment plays a role in language evolution. Data on
33 actual language change, coming from historical linguistics and sociolinguistics, is currently
34 being used to constrain the repair and consolidation strategies of agents in grammatical language
35 games and data from cognitive linguistics and particularly cognitive semantics is used to
36 constrain the range of possible conceptualizations that could be the target of experiments. The
37 theoretical tools developed in statistical physics and complex systems science have recently
38 acquired a central role for the study of Language Games. The suite of methods developed in
39 these fields has indeed allowed to address quantitatively such issues as the scaling of relevant
40 features of the models with the system size (e.g. convergence time or memory requirements
41 (Baronchelli et al., 2006a, 2008), the impact of different underlying topology on global behaviors
42 (e.g. homogeneous mixing (Baronchelli et al., 2008) vs. regular lattices (Baronchelli et al.,
43 2006b) vs. complex networks (Dall'Asta et al., 2006a, 2006b)), and the detailed study of

1 convergence dynamics (Baronchelli et al., 2008). Thus, for example, it has been shown that
2 complex networks are able to yield, at the same time, the fast convergence observed in
3 unstructured populations and the finite memory requirements of low dimensional lattices
4 (Dall’Asta et al., 2006a, 2006b). Moreover, agents’ architectures and interaction rules have been
5 significantly simplified to allow thorough analysis, and this has allowed to pinpoint the crucial
6 ingredients responsible for the desired global co-ordination. The pursuit of simplicity, along with
7 the novelty of the complex systems approach to this field, has so far limited the investigations
8 mostly to the study of the Naming Game and of the Category Game (in which the population
9 ends up with a shared repertoire of categories) (Puglisi et al, 2008). Research is however ongoing
10 in order to tackle higher order problems, such as the emergence of compositionality (De Beule,
11 2008). Experiments with human subjects show that humans can evolve communication systems
12 although some are better than others, mostly because of differences in social attitudes. Of course
13 the greatest challenge is to scale these experiments up to the level of grammatical languages.
14 Recent examples already showing the formation of case grammars, tense-aspect-mood systems,
15 or determiner systems lead to optimism (see e.g. Van Trijp, 2008).

16 *Genetic Evolution*

17 A third class of models explores the role of biology by modeling the genetic transmission of
18 language. Agents are created based on a model of a genome that codes directly the lexicon or
19 grammar of their language. Agents then engage in interactions that determine their fitness, and
20 based on communicative success they have a higher chance to reproduce in the next generation.
21 Due to random mutations and crossover, offspring has slightly different genomes, possibly
22 giving higher communicative fitness which then leads to further propagation. These models use
23 very similar techniques as those used in genetic algorithms and they sprang up first in the context
24 of Artificial Life (see Cangelosi & Parisi, 1998). Given that the explicit genetic coding of lexicon
25 and grammar is highly implausible from a biological point of view, more recent models have
26 considerably weakened this assumption, and encode only strong biases and universal constraints
27 on possible languages. This is particularly the case for the ENGA model (Szathmary, 2007).
28 ENGA is an ambitious framework that covers not only the genetics but also the neuro-
29 developmental processes in a biologically realistic way. Linguistic inventories are not coded
30 genetically but acquired by a learning process. The ENGA model therefore attempts to cover the
31 whole ground from genetic to developmental and learning processes.

32 **1.2. Aggregate models**

33 In addition to agent-based models, there is extensive research to construct macroscopic models
34 of language evolution and language dynamics.

35 *Game Theoretic Models of Language Evolution*

36 The main paradigm being explored draws from the tradition of evolutionary game theory in
37 order to focus on the role of imitation in cultural transmission. Imitation (or re-use) applies both
38 to the adaptation of linguistic performance between adult speakers and the acquisition of
39 language by infants. Imitation is framed as a form of replication. An evolutionary dynamics
40 ensues in any population of replicating entities, provided the entities in the population vary in
41 certain heritable characteristics, and replicative success is correlated with this variation. This is a
42 crucial difference to the iterated learning paradigm, where every individual grammar participates

1 equally in language replication. On the other hand, the game theoretic model – as a form of a
2 selectionist model – assumes faithful replication, while replication under iterated learning may
3 be imperfect. Under certain simplifying assumptions – like the postulation of an infinite
4 population and a continuous time – such an evolutionary dynamics can be described by a system
5 of ordinary differential equations. In language evolution this dynamics is necessarily nonlinear
6 because selection is frequency dependent. This can, for instance, be illustrated by the
7 development of vocabulary: whether a candidate for a neologism catches on in a linguistic
8 community (i.e. becomes replicated) depends on whether or not there already is another word for
9 the same concept within this linguistic community. This indicates that the overall frequency
10 distribution of words is a decisive factor for the fitness of each individual word. A similar point
11 can be made for other linguistic units, ranging from phonemes to syntactic constructions.

12 Frequency dependent selection can be modeled by means of replicator dynamics within the
13 mathematical framework of *evolutionary game theory* (Maynard Smith 1982, Hofbauer &
14 Sigmund 1998). A model of a communication game consists, in its simplest incarnation, of

- 15 • a space of meanings and a space of forms,
- 16 • a space of production grammars (mappings from meanings to forms),
- 17 • a space of comprehension grammars (mappings from forms to meanings), and
- 18 • a utility function, i.e. a measure of success for a pairing of grammars, depending on the
19 success of communication and complexity of the grammars involved.

20 Further parameters may be added, like a biased *a priori* probability distribution over meanings,
21 or a confusion matrix for noisy transmissions of forms.

22 There are several off-the-shelf theorems from biomathematics regarding stability conditions for
23 evolutionary games. Such theorems sometimes render it straightforward to identify the attractor
24 states of the replicator dynamics without actually delving into the complexities of the underlying
25 nonlinear differential equations.

26 The biomathematics literature contains a variety of results concerning the evolution of
27 communication, where strategies (“grammars”) are assumed to be innate and replication is
28 interpreted in the biological sense (e.g., Wärneryd 1993, Trapa & Nowak 2000, Nowak &
29 Krakauer 1999, Nowak, Krakauer & Dress 1999, and Jäger 2008a). These authors mainly
30 consider biological evolution, and they assume that communicative success is correlated with
31 biological fitness, i.e. the number of fertile offspring. However, their results are general enough
32 that they can be extrapolated to cultural evolution. The background assumption here is that
33 communicative success of a certain behavioral trait is positively correlated with its likelihood to
34 be imitated, i.e. its cultural fitness. Possible applications of evolutionary game theory to the
35 study of the cultural evolution of language (in the sense described above) are investigated in a
36 series of papers by Gerhard Jäger and Robert van Rooij (Jäger 2007, 2008b, Jäger & van Rooij
37 2007).

38 Game-theoretic research in language evolution has suggested a formal framework which is quite
39 useful within this paradigm. “Universal grammar” or a pre-existing bias of grammar learning
40 can be represented in the following abstract manner. Suppose we have a finite alphabet (a finite
41 set of symbols) (see Nowak et al 2001, Komarova et al 2001, Komarova & Nowak 2001, 2003,

1 Nowak & Komarova 2001). A language is a probability distribution defined on a set of strings
2 composed of the symbols of the alphabet. The allowed languages can be represented as
3 probability distributions on a collection of (intersecting) sets. Then a learning mechanism is a
4 way to “navigate” in this collection of sets. Pair-wise similarity among languages can be
5 expressed as a matrix. The process of learning then is a sequence of hypotheses of a learner in
6 response to the input of a teacher (or teachers), which is a number of strings compatible with the
7 teacher(s)’ grammar. This framework allows one to use the machinery from mathematical
8 learning theory, and connect natural language evolution with insights from computer
9 science/machine learning.

10 **1.5 Summary**

11 There are obvious relations, complementarities and continuities between these approaches and
12 paradigms. The game-theoretic paradigm focuses on the selectionist dynamics of the language
13 itself, whereas language game models use an agent based approach, focusing on the cognitive
14 mechanisms by which agents use, invent and coordinate language so that the selectionist
15 dynamics of language emerges. The Iterated Learning paradigm focuses on the role of bias and
16 the vertical transmission bottleneck and therefore tends not to integrate the issue of
17 communicative success, cognitive effort or population dynamics into the models, whereas the
18 Language Game paradigm considers vertical transmission as an additional but not crucial effect
19 on language evolution. Pursuing these different approaches provides the opportunity to explore
20 how different factors such as learning, communication, and population structure influence the
21 process of language evolution.

24 **2. Linguistic Representations and Processes**

25 Given that this Forum was focused on syntax, it is relevant to ask the question what kind of
26 representations for grammar are being used in language evolution models and what kind of
27 syntactic operations and grammatical processes have been incorporated into these models. It
28 turns out that researchers working on iterated learning and game-theoretic approaches generally
29 try to use existing ‘symbolic’ formalisms or neural network models. Some have argued however
30 that the requirements of evolvability put additional constraints on the nature of grammatical
31 representations and processing and this has led to some work on novel grammar formalisms
32 which can cope with emergent grammar.

33 *Symbolic Grammars*

34 There are a variety of grammatical formalisms in the theoretical linguistics literature, some of
35 which have been utilized in evolutionary models whereas others, such as minimalism (Chomsky
36 1995), have not (possibly because they are less easily embedded in theories of processing).
37 Examples of formalisms which have been deployed with minimal modification include
38 Optimality Theory (Jäger 2004), Extended Categorical Grammar (Briscoe 2000) and Context Free
39 Grammars (Zuidema 2002). All such models require the embedding of the formalism into a
40 theory of grammar learning and processing. Modelers have drawn on existing proposals from the
41 literature, such as Bayesian parameter estimation, compression based algorithms, or non-

1 statistical parameter setting algorithms for implementing the learning mechanisms used in
2 vertical transmission (see Griffiths and Kalish 2007, Briscoe 2000).

3 *Simple Recurrent Networks*

4 Other language evolution models have avoided the explicit representation of hierarchical
5 structures, syntactic and semantic categories and grammatical rules, deploying distributed and
6 subsymbolic representation. A popular alternative is Simple Recurrent Networks (SRNs, Elman,
7 1990). In SRNs, knowledge of language is learnt from the presentation of multiple examples
8 from which the networks learn to process syntactic structure. The general aim of such models is
9 to capture observable language performance, rather than idealized linguistic competence
10 (Christiansen, 1992; Christiansen & Chater, 1999). Much of this work has an emphasis on the
11 integration of multiple sources of probabilistic information available in the input to the
12 learner/speaker/hearer (e.g., from the perceptuo-motor system, cognition, socio-pragmatics, and
13 thought as discussed in the chapter by Kirby, Christiansen & Chater, this volume). Although
14 much of this work tends to target small fragments of language for the purpose of close modeling
15 of psycholinguistic results (e.g., Christiansen & Chater, 1999; MacDonald & Christiansen,
16 2002), some efforts have gone into scaling up models to deal with more realistic language
17 samples, such as full-blown child-directed speech (Reali, Christiansen & Monaghan, 2003). In
18 this framework grammatical processing can be conceptualized as a trajectory through a high-
19 dimensional state-space afforded by the hidden unit activations of the network (e.g., Elman,
20 1990), potentially suggesting an alternative perspective on constituency and recursion in
21 language (Christiansen & Chater, 2003). These models do not include explicit grammar
22 formalisms but the behavior of the networks can in some cases be described in terms of such
23 formalisms.

24 *Formalisms designed for grammar evolution*

25 Some researchers have been developing novel formalisms to be used specifically in language
26 game experiments. This is particularly the case for Fluid Construction Grammar (FCG). FCG (de
27 Beule & Steels, 2005) uses representational mechanisms already employed in several existing
28 symbolic grammar formalisms like HPSG (Sag, et.al. 2006) or Lexical-Functional grammar
29 (Kaplan & Bresnan, 1982) such as a feature-structure based representation of intermediary
30 structures during parsing and production, a constraint-based representation of linguistic rules so
31 that they can be applied in a bi-directional fashion, and unification-style mechanisms for the
32 application of these rules. FCG is in line with other construction grammar formalisms (such as
33 Embodied Construction Grammar, Bergen & Chang, 2004) in the sense of supporting the
34 explicit representation and processing of constructions, which is de-emphasised in Minimalism.
35 But FCG on the other hand has various additional facilities to enable language evolution
36 experiments: (i) Individual agents represent a multitude of hypotheses about the emerging
37 language, and are therefore able to handle variation in language use, (ii) rule application is
38 flexible allowing the violation of constraints and robust parsing and production so that sentences
39 can be understood even if they are not entirely grammatical (according to the preferred grammar
40 of the agent), (iii) the different variants compete within the individual when it has to make
41 decisions about how to express something or interpret something and, as an emergent effect,
42 within the population for dominance in the emergent language, and (iv) rather than coding
43 systematicity in terms of more abstract rules, FCG maintains links between the rules, based on

1 how the rules are formed through composition of other rules. These links are then used for
2 assigning credit or blame after a game, allowing the implementation of a multi-level selectionist
3 dynamics. Due to these features, FCG exhibits dynamical systems properties seen in network
4 representation systems of grammar that do not rely on symbolic structures, such as connectionist
5 networks or recurrent neural networks, while at the same time incorporating many ideas from
6 decades of research into theoretical and computational linguistics.

7 **2.3. Summary**

8 The computer simulations carried out in evolution of language research rest on a variety of
9 formalisms to represent the inventory of the lexicon and grammar of the emerging language in
10 the first place. In choosing a particular formalism, the researcher makes a commitment to what
11 aspects of language are isolated for an inspection of their role in evolutionary dynamics, and
12 what others are (implicitly) excluded.

14 **3. How can Modeling (Already) Inform the Study of Language Evolution?**

15 Although the mathematical and computational modeling of language evolution is still in its
16 infancy, there are already quite a few results that show the power of the approach and that may
17 be of interest to biologists and linguistics.

18 **3.1. Two main sources of added insight**

19 Computational modeling, like in any other field, enables two powerful avenues for accruing
20 scientific insight:

21 **Formal analysis.** Computational models have to be rigorously formalized to make them
22 operational on computers. When a simulation is running, all aspects of the simulation can be
23 recorded, including the population aspects. The same is true in robotic experiments where all
24 perceptual states, motor states, and the full details of all processes going into language
25 production and understanding can be tracked, something not possible with human subjects. This
26 full access to relevant data makes the models amenable to mathematical analyses. Typical
27 questions that can be answered by the analytical methods provided by nonlinear dynamics, game
28 theory and statistical physics concern asymptotic properties of evolutionary dynamics, the
29 dependence of these dynamics on scaling parameters, or the prediction of sudden and dramatic
30 changes (phase transitions).

31 **Simulation studies.** Carrying out simulations on a computer differs from carrying out real-life
32 experiments in two crucial respects. First, the simulated piece of reality is *completely* specified.
33 Second, one has *full control* over varying experimental conditions. There are risks and
34 opportunities under these circumstances. An obvious pitfall is that the simulation may miss a
35 crucial component of the real-life target system – this is the problem of abstraction. However, it
36 should be noted that, in principle, in experimental designs involving human subjects the same
37 problem is present: a particular experimental design may prevent real-life-relevant mechanisms
38 from taking effect. The benefits added to empirical experiments (which remain indispensable) by
39 simulation studies are, in our view, the following:

- 1 • A systematic *exploration of large hypothesis spaces* is made possible due to the speed
2 and low cost of simulations. This both facilitates the generation of new scientific
3 hypotheses, and the testing of existing ones.
- 4 • Model simulations can give *existence proofs* for the efficiency of certain mechanisms to
5 achieve a certain effect – always, of course, modulo the modelling assumptions.
- 6 • In a related vein, model simulations can give *non-uniqueness proofs* if the same ultimate
7 effect can be obtained by different mechanisms. Such demonstrations are helpful in
8 precluding an early "contraction" to a single explanatory venue in theory development.
- 9 • Simulations are *replicable* across different laboratories by sharing code.
- 10 • Critiquing and improving simulation setups is *transparent*, because it is explicit how
11 assumptions become operationalized in the designs.

12 If one is carefully conscious of the assumptions that go into a simulation model, research based
13 on such models can decidedly "open up" the space of possible theories in a field, raising the
14 awareness of alternative theories. To demonstrate this point, in this section we present a number
15 of examples that have arisen from the work of group members and which were discussed at the
16 meeting.

17 **3.2. Examples**

18 In the following we list some of the important contributions to understanding language
19 evolution, derived from work carried out by authors of this chapter.

20 **Magnification of Learning Bias through Cultural Transmission**

21 Mathematical analyses of the iterated learning model described above provides some interesting
22 insights into the relationship between the inductive biases of language learners -- the factors that
23 lead them to find it easier to learn one language than another, as might be the consequence of
24 genetic constraints on language learning -- and the kinds of languages that will be spoken in a
25 community. As discussed by Kirby, Christiansen, and Chater (this volume) and Briscoe (this
26 volume), one way to capture the inductive biases of learners is to assume that they identify a
27 language from a set of utterances by applying Bayesian inference, with a "prior" distribution
28 encoding which languages learners consider more probable before seeing any data. Languages
29 with higher prior probability can be learned from less evidence, and the prior thus reflects the
30 inductive biases of the learner. Analyses of iterated learning with Bayesian agents show that the
31 relationship between the prior and the languages that are ultimately produced via cultural
32 transmission can be complex (Griffiths & Kalish, 2007; Kirby, Dowman, & Griffiths, 2007).
33 Specifically, iterated learning can magnify weak inductive biases, with a slight difference in the
34 prior probabilities of two languages resulting in a significant difference in the probability of
35 those languages being produced via cultural transmission. These mathematical results suggest
36 that strong genetically-encoded constraints on learning may not be necessary in order to explain
37 the structure of human languages, with cultural evolution taking on part of the role that might
38 otherwise have been played by biological evolution.

39 **Restricting the Space of Possible Grammars**

40 It is tempting to reconstruct the notion of a linguistic universal as a property that every language
41 with a grammar that can be cognitively represented and learned by humans – i.e. a language that
42 conforms to “Universal Grammar” in the Chomskyan sense – shares. Evolutionary models

1 indicate that there may be other sources of universals. Briefly put, a *possible language* must also
2 be attainable under the evolutionary dynamics of language transmission.

3 In Jäger (2004), this basic idea is illustrated with a particular implementation. According to
4 *Optimality Theory*, Universal Grammar defines a finite set of constraints, and each particular
5 grammar is characterized by a linear ordering of these constraints. To account for certain strong
6 typological tendencies, Aissen (2003) proposed to restrict the space of possible grammars further
7 by imposing certain sub-hierarchies of constraints that are never violated.

8 Following proposals by Boersma (1998), Jäger implements a stochastic learning algorithm for
9 optimality theoretic grammars. However, unlike Boersma, Jäger assumes that language
10 acquisition is bidirectional, i.e. the learner tries both to mimic the production behaviour and the
11 comprehension behaviour of the teacher. It turned out that some constraint rankings are strictly
12 not learnable at all. Among the remaining space of learnable grammars, some are more robustly
13 learnable than others. After iterating the learning procedure a few dozen or hundred of times
14 (where in each generation, the former learner becomes the teacher and produces utterances on
15 the basis of his acquired grammar), only constraint rankings that conform to Aissen's prediction
16 were observed.

17 **The Co-Evolution of Categories and Names**

18 One of the big debates in language studies concerns the question of how far perceptually
19 grounded categories, such as colors, influence and are influenced by language that expresses
20 these categories. From a Whorfian point of view there is a strong interaction whereas those
21 arguing for strong modularity have argued that categories are innate or induced from empirical
22 data and language are just labels for existing categories. Although color categorization and color
23 naming does not relate directly to grammar, we include this theme here because it exemplifies
24 the quality of insight that can be obtained from modelling studies, and because categorization
25 and naming are prerequisites for grammatical language. Research on using for language games
26 for studying the co-evolution of categories and names started with the BBS paper by Steels and
27 Belpaeme (2004) in which agent-based models of color naming and categorization were
28 developed and systematically compared. This paper showed that although a genetic evolution of
29 color categories was possible, it not only took a long time, but also did not lead to a system that
30 was adaptive, and did surely not lead to universal categories unless populations remained
31 homogeneous. The paper also showed that a purely learning-based approach did not lead to an
32 explanation for trends in color categories and neither to sufficient coherence in a population to
33 explain how a successful communication was possible. More recently this research was extended
34 in two directions.

35 *Deepening the Complex Systems Approach to Color Categorization*

36 The Category Game (Puglisi et al., 2008) is a language game that aims at describing how a
37 population of agents can bootstrap a shared repertoire of linguistic categories out of pairwise
38 interactions and without any central coordination. The prototypical example of the phenomena
39 the model addresses is given by color categorization. Individuals may in principle perceive
40 colors in different ways, but they need to align their linguistic ontologies in order to understand
41 each others. In the game, a population of N individuals is committed to the categorization of a
42 single analogical truly-continuous perceptual channel, each stimulus (or "object") being a real

1 number in the interval $[0,1)$. A categorization is identified with a partition of the interval $[0,1)$ in
2 discrete sub-intervals, or perceptual categories. Individuals have dynamical inventories of form-
3 meaning associations linking perceptual categories to words representing their linguistic
4 counterparts, and they evolve through elementary language games. At the beginning all
5 individuals have only the trivial perceptual category $[0,1)$. At each time step two individuals are
6 selected and a scene of M stimuli is presented. The speaker discriminates the scene, if necessary
7 refining its perceptual categorization, and names one object. The hearer tries to guess the named
8 object, and based on her success or failure, both individuals rearrange their form-meaning
9 inventories. The only parameter is the just noticeable difference (JND) of the individuals. The
10 probability distribution from which stimuli are randomly chosen, finally, characterizes the kind
11 of simulated environment.

12 The main result is the emergence of a shared linguistic layer in which perceptual categories are
13 grouped together into emerging linguistic categories to guarantee communicative success.
14 Indeed, while perceptual categories are poorly aligned between individuals, the boundaries of the
15 linguistic categories emerge as a self-organized property of the whole population and are
16 therefore almost perfectly harmonized at a global level. Interestingly, the model reproduces a
17 typical feature of natural languages: despite a very high resolution power and large population
18 sizes, the number of linguistic categories is finite and small. Moreover, a population of
19 individuals reacts to a given environment by refining the linguistic partitioning of the most
20 stimulated regions, while non-uniform JNDs (like for instance the human JND function relative
21 to hue perception) constrain to some extent the structure of the emergent ontology of linguistic
22 categories.

23 *The Evolutionary Game Theory Approach to Color Categorization*

24 The following simple framework has been designed in order to investigate the influence of
25 various realistic features (linguistic, psychological and physiological) on the shared color
26 categorization (see Komarova et al 2007, Komarova & Jameson 2008). The space of colors is
27 represented as a 3-D spheroid, or a lower-dimension subset of that. Color categorization of an
28 agent is modeled as a stochastic matrix, which specifies the probability of color names used to
29 denote color exemplars, which are the psychological representations of, say, various color hues.
30 The process of color categorization is simulated as repeated discrimination and communication
31 games played by a number of agents. The discrimination game consists of two exemplars
32 presented to an agent, followed by the agent using his categorization matrix to assign color terms
33 to the two exemplars. The outcome of the game (success or failure) is decided based upon the
34 following pragmatic criterion. If the two color exemplars are “close” to each other and are
35 classified as the same category, then the game is a success; if they are classified differently, then
36 it is a failure. On the other hand, if the exemplars are “far apart”, then for success they have to be
37 categorized as different. The measure of “closeness” of two exemplars is specified (in the
38 simplest case) by a single “pragmatic similarity” parameter. After each run of the discrimination
39 game, the categorization matrix of the agent is modified to strengthen the more successful
40 category and weaken the less successful one. Communication between agents is modeled by
41 pairs of agents playing the discrimination game and the less successful agent modifying its
42 categorization matrix accordingly. As a result of a number of iterations of this game, a
43 population of agents arrives at a shared categorization system, which possesses the following

1 qualities: (i) the exemplar space is equipartitioned into a (predictable) number of distinct,
2 deterministic color categories, (ii) the size of color categories is uniquely defined by the
3 pragmatic similarity parameter, (iii) the location of category boundaries possesses rotational
4 symmetry. While this skeleton produces results reasonable from a psychological point of view,
5 the main *raison d'être* for this model is to investigate various realistic constraints on color
6 categorization. For example, non-uniformities in the “color diet” lead to differential
7 convergence rates of different color categories. The inhomogenities of exemplar space (the non-
8 uniformity of the pragmatic similarity parameter) lead to changes in size and number of color
9 categories. Finally, inhomogenities in the agent population can also change the structure of the
10 common categorization system. Interestingly, the presence of even a small number of abnormal
11 observers (e.g. dichromats) in the population leads to the anchoring of color boundaries to a
12 subset of possible locations. These locations are defined by the confusion regions in the
13 dichromats' color representation. Empirical data of confusion spectra of abnormal color
14 observers can be incorporated to generate specific color boundary predictions and to deduce
15 how the color categorization of various populations is influenced by the population
16 inhomogeneities (see Jameson & Komarova, 2009).

17 **The Emergence of Linguistic Ontologies**

18 The final example that shows how modeling can lead to the opening up of new theoretical
19 avenues and ideas in language evolution comes from the domain of grammar. Grammar exploits
20 syntactic devices (such as word order or morphology) in order to express additional aspects of
21 meaning, such as discourse structure, thematic relations (predicate-argument structure), tense-
22 aspect-mood, determination, scoping constraints on anaphora, etc. In all linguistic theories of
23 today the rules of grammar are expressed using an ontology of syntactic and semantic categories.
24 These syntactic categories include parts of speech (e.g., noun, verb, adverb), types of
25 constituents (e.g., noun phrase, relative clause), syntactic constraints (e.g., agreement,
26 precedence), syntactic features (e.g., nominative, masculine, neuter), etc. The semantic
27 categories include categorisations of temporal aspects in terms of tense, aspect, or mood,
28 semantic roles such as agent or beneficiary, categories used for conceptualising discourse, like
29 topic/comment, different shades of determination (e.g. definite/indefinite, count/mass),
30 classifiers (as used in Bantu languages), deictic references both for use inside and outside
31 discourse, epistemic distinctions, and so on. A complex grammar undoubtedly requires
32 hundreds of such categories. A fundamental question in understanding the origins and evolution
33 of language is therefore where such a linguistic ontology might be coming from.

34 There is a common (usually hidden) assumption among many theorists that linguistic ontology is
35 universal and innate, but that does not explain yet how it originates. Typologists have argued that
36 linguistic categories are to a large extent language-dependent (Haspelmath, 2007) and historical
37 linguists have shown that categories change over time (Heine & Kuteva, 2008). This suggests
38 that linguistic categories may be similar to categories in other domains of cognition (such as the
39 color categories discussed earlier), in the sense that they are culturally constructed and
40 coordinated.

41 Recent language game experiments in the formation of a case grammar (see Steels, this volume)
42 have shown that the formation of linguistic ontologies is entirely possible. Concretely, semantic
43 roles as needed in case grammar have been shown to arise when agents are trying to reuse by

1 analogy semantic frames that have already been expressed in the emergent language. This reuse
2 becomes licensed when particular predicate-argument relations are categorised in the same way
3 as those already used in the existing semantic frames. Progressively, semantic roles get thus
4 established and refined, partially driven by the semantic analogies that make sense in the real
5 world domain that generates the topics in the language game and partly by the conventions that
6 are being enforced by the emergent language (Van Trijp, 2008).

7 **3.3. Summary**

8 The examples discussed in this section illustrate some of the ways in which models of the
9 cultural evolution of language can contribute to our understanding of its origins. By identifying
10 what aspects of the properties of languages can be produced by cultural evolution alone, these
11 models remove some of the explanatory burden from biological evolution, providing a more
12 realistic target for research into the origins of language. In broad terms, these models illustrate
13 how learning, communication, and population structure affect the languages that emerge from
14 cultural evolution, providing potential explanations for two of the most important aspects of
15 human languages: their consistent properties across communities – language universals – and the
16 coherence of linguistic systems within communities. In iterated learning models, universals
17 emerge as the result of learning biases or the goals of communication, and coherence is the result
18 of the strength of these biases and the structure of the interactions with other individuals. In
19 language game experiments, universal trends emerge due to constraints coming from
20 embodiment, the cognitive mechanisms recruited for language, the challenge of communication,
21 and the selectionist dynamics that emerges in populations of adaptive communicating agents.
22 While there are still many questions to explore, these basic results help to illustrate the kinds of
23 forces that influence the structure of human languages.

24 25 **4. Suggestions for Future Research**

26 Given that there is a broad variety of paradigms and modeling efforts, there are also many
27 possible avenues for deepening current results or for exploring new avenues of research. This
28 section describes a number of suggestions without any claim to be exhaustive. Generally
29 speaking, there are also many possible avenues for deepening current results or for exploring
30 new avenues of research. Generally speaking, we can expect models to be developed that focus
31 on quite different aspects of language evolution and that will be formulated at very different
32 levels of abstraction. It will be important to establish the relationships between these models,
33 such as identifying to what extent a simpler and more abstract model can be understood as an
34 approximation to a more elaborate one.

35 **Toward a Tighter Coupling Between Models and Laboratory Experiments**

36 An important direction for future research is developing a tighter coupling between models and
37 laboratory experiments. There are two ways in which conducting laboratory experiments in
38 cultural evolution can complement the insights provided by mathematical and computational
39 models. First, they provide a direct way of testing the predictions of these models, allowing us to
40 ensure that the claims that we make about cultural evolution are actually borne out when these
41 processes involve real people rather than abstract agents. For example, Kalish, Griffiths, and

1 Lewandowsky (2007) and Griffiths, Christian, and Kalish (2008) have conducted direct tests of
2 the key prediction that arises from models of iterated learning with Bayesian agents – that
3 structures that are easier to learn will be favored by the process of cultural transmission – by
4 conducting laboratory experiments in which the structures transmitted by iterated learning were
5 categories and functional relationships between variables for which previous research in
6 cognitive psychology had established results on difficulty of learning. However, laboratory
7 experiments can also be valuable for a second reason: they provide us with a closer
8 approximation to the true processes involved in language evolution. The models discussed earlier
9 make assumptions both about how information is passed between agents, and the learning
10 mechanisms used by those agents. Conducting laboratory experiments in which information is
11 passed between agents in the way described by a model, but where the agents are real human
12 beings, removes one level of approximation from these models, allowing us to explore the
13 plausibility of processes of cultural transmission as an account of why languages have the
14 properties they do (Dowman, Xu, & Griffiths, 2008). The experiment described by Kirby,
15 Christiansen, and Chater (this volume) is of this kind, showing that iterated learning with human
16 learners produces compositional structures. Further experiments testing models of language
17 evolution and evaluating the impact of different forms of cultural transmission can help us
18 develop models that provide a closer match to human behavior, and to assess the contributions of
19 different kinds of evolutionary forces.

20 **Toward a Tighter Coupling Between Models and Data from historical linguistics**

21 Much is known about the historical evolution of human languages over the past 5000 years. This
22 research shows that there are recurrent patterns of grammaticalisation and lexical change and
23 detailed case studies exist how a language has developed determiners, or a case system, or a
24 tonal system, etc. (see e.g. Heine & Kuteva, 2008). It is therefore obvious that these results
25 should constrain models of language evolution. Although it will of course never be possible to
26 reconstruct the actual evolution of human languages, it might be possible to see similar
27 grammaticalisation phenomena as in human languages.

28 **Modeling the Potential Role of Exaptation on Language Evolution**

29 It is widely assumed that language in some form or other originated by piggybacking on pre-
30 existing mechanism – exaptations – not dedicated to language. A possible avenue of language
31 evolution modeling involves testing the possible effects for language evolution of particular
32 hypothesized exaptations. For example, improved sequential learning of hierarchically organized
33 structure in the human lineage has been proposed as a possible preadaptation for language
34 (Christiansen & Chater, in press; Conway & Christiansen, 2001), in part based on work in
35 language acquisition (Gómez & Gerken, 2000) and genetic data regarding the potential role of
36 *FOXP2* in sequential learning (discussed elsewhere in this volume). Reali & Christiansen (in
37 press) have explored the implications of such assumptions by determining the effect of
38 constraints derived from an earlier evolved mechanism for sequential learning on the interaction
39 between biological and linguistic adaptation across generations of language learners. SRNs were
40 initially allowed to evolve “biologically” to improve their sequential learning abilities, after
41 which language was introduced into the population, comparing the relative contribution of
42 biological and linguistic adaptation by allowing both networks and language to change over
43 time. Reali & Christiansen’s (in press) simulation results supported two main conclusions: First,

1 over generations, a consistent head-ordering emerged due to linguistic adaptation. This is
2 consistent with previous studies suggesting that some apparently arbitrary aspects of linguistic
3 structure may arise from cognitive constraints on sequential learning. Second, when networks
4 were selected to maintain a good level of performance on the sequential learning task, language
5 learnability is significantly improved by linguistic adaptation but not by biological adaptation.
6 Indeed, the pressure toward maintaining a high level of sequential learning performance
7 prevented biological assimilation of linguistic-specific knowledge from occurring. Similarly, it
8 may be possible to investigate the potential effects of other hypothesized exaptations on the
9 relative contribution of cultural evolution and genetic assimilation to language evolution.

10 In the same line, several language game experiments have examined how generic cognitive
11 mechanisms could become recruited for language, pushed by the needs to solve specific
12 problems in communication or in bootstrapping an efficient system (Steels, 2007). For example,
13 perspective and perspective reversal is often lexicalized in human languages in order to avoid
14 ambiguity from which point of view a spatial relation should be interpreted.

15 **Effects of Biased Unfaithful Copying**

16 When empirical predictions are derived from dynamical models, the notion of an *equilibrium* is
17 central. In the evolutionary context, we expect systems to spend most of their time in an
18 *evolutionarily stable state*. The insights from historical linguistics, especially regarding
19 *grammaticalization*, indicate that language never actually reaches such a stable state. (This
20 statement might be too bold in its generality. Some aspects of language are certainly in
21 equilibrium most of the time. A good example might be vowel systems.) Rather, languages
22 perpetually change in a partially predictable way. Complex morphology tends to be reduced over
23 time and to disappear altogether eventually. An example is the loss of case distinctions from
24 Latin (five cases) to French (no case distinctions). On the other hand, lexical morphemes are
25 recruited to serve grammatical functions. A recent example is the use of the item “going to” in
26 contemporary English to express future. This recruitment usually concurs with phonological
27 reduction, like the change from “going to” to “gonna”. Grammatical words tend to get further
28 reduced to affixes – an example would be the regular German past tense morpheme “t” that is
29 originally derived from the Germanic verb for “do”.

30 The macroscopic consequence of these processes is that languages continually change their
31 grammatical type, moving from synthetic to analytic due to reduction of morphology, and back
32 to synthetic due to recruitment of lexical items for grammatical purposes and their subsequent
33 reduction to affixes. The underlying microdynamics involves biased unfaithful copying – words
34 and phrases are not imitated verbatim but phonetically reduced and semantically modified. The
35 challenge for evolutionary models is to connect these two aspects in such a way that the
36 directedness of language change is connected to empirical insight about unfaithful replication in
37 language use. Deutscher (2005) in his book proposes a verbal model which resembles the
38 sociolinguistic arguments of Labov (2001). Individuals often innovate new speech forms in an
39 effort to find a more emphatic or colorful way of phrasing an idea or grammatical function.
40 Conventional forms bore us while prose or speech stylists that play with the limits of convention
41 attract attention. When prestigious people do this, the new speech form tends to spread.
42 Sometimes the motivations for innovation are social; people seem to favor forms of speaking
43 that differentiate them from social others. In other words, linguistic equilibria are weakly

1 constrained in that communicating individuals must agree *sufficiently* on meanings for
2 communication to be possible. But a speech community can easily cope with a modest rate of
3 innovation driven by social and aesthetic forces. To our knowledge, these mechanisms have not
4 been incorporated into formal models except in the special case of symbolic markers of group
5 boundaries (McElreath et al. 2003)

6 **Long-Term Language Change Dynamics: a Mathematical Perspective**

7 It appears that language modeling poses challenges for the existing mathematical methods
8 commonly used to describe emerging and dynamical real-life phenomenon. A ready example
9 comes from language games. Language game solutions may vary with regards to their stability
10 properties depending on the type/purpose of the model use, and depending on the exact question
11 we address. In certain situations interesting quasi-stable solutions are attained. One instantiation
12 comes from modeling color categorization in people, where the shared population categorization
13 solution cannot be described as a stable solution of a dynamical system, or a stationary
14 probability distribution of a stochastic process. In the Category Game (Puglisi et al, 2008), even
15 though the only absorbing state is the trivial one in which all the agents share the same unique
16 word for all their perceptual categories, there are clear signatures of a saturation with time of
17 metastable states with a finite and “small” number of linguistic categories. This observation
18 suggests an analogy with glassy systems in physics (Mezard et al 1987), and this view is
19 confirmed also by quantitative observations. Thus, in this framework interesting solutions would
20 be long lived (strongly, e.g. exponentially, dependent on the population size) pre-asymptotic
21 states. In other models of color categorization, the shared population categorization solution
22 appears dynamically stable on a certain time-scale, but it may drift or cycle (while retaining
23 global topological structure) on longer time-scales, depending on the particular constraints (see
24 Komarova et al 2007). Mathematical properties of such solutions have not been investigated in
25 detail but their understanding is important because conventional methods do not grasp the
26 relevant properties of such solutions. The application of new mathematical technologies thus
27 developed will be wide, as it has implications in the dynamics of populations of learners trying to
28 achieve shared solutions on (possibly very complex) topological semantic spaces.

29 **Selectively Neutral Mechanisms of Linguistic Evolution**

30 A further direction for future research is understanding to what extent processes of selection are
31 necessary in order to explain the properties of languages. In biology, selectively neutral
32 processes such as mutation and genetic drift have been identified as playing a significant role in
33 accounting for genetic variation (Kimura, 1983). It remains to be seen whether linguistic
34 variation is best analyzed as the result of selective pressures acting on the properties of
35 languages, or the outcome of selectively neutral processes that are the cultural equivalents of
36 mutation and drift. Answering this question requires developing a "neutral theory" for language
37 evolution. In this case, the analogue of mutation is the variation that is produced as a
38 consequence of failed transmission of languages through the "learning bottleneck" produced by
39 the fact that learners only observe a finite number of utterances. Iterated learning models thus
40 provide a starting point for developing a neutral theory, and understanding which properties of
41 languages can be produced by iterated learning and which properties cannot thus constitutes an
42 interesting direction for future research.

1 **Replicators**

2 This discussion opens up the question whether or not we should be thinking about cultural
3 transmission/social interaction models in terms of competition among replications, or more
4 excitingly, in terms of different levels or types of replication. On the one hand, it is tempting to
5 propose (as outlined in Kirby, 2006) that the emergence of syntax marks a change from one type
6 of replicator (solitary replicators) to another (ensemble replicators), to use terms from Szathmary
7 (2000).

8 On the other hand, this gives rise to the question of what exactly these replicators are, and
9 whether their dynamics are best described in terms of selection at all. It appears that the answers
10 to these questions vary enormously depending on one's perspective on the best way to represent
11 the knowledge being acquired/adapted by individuals and the mechanisms for acquiring that
12 knowledge. For example, one view of language might propose that we internalise a set of
13 constructions (e.g. Croft, 2000) that have a fairly straightforward relationship with utterances. In
14 this view, we might reasonably think of these constructions as replicators, with selection being
15 driven by speakers choosing among constructions to use to produce an utterance. Or perhaps we
16 could think of learners as providing selection pressure, with the constructions that produce the
17 most evidence for their existence in the data available to the learner ending up being the most
18 stable through the learning bottleneck. Here, we can imagine constructions competing for place
19 in the learners' input.

20 Another view might be that a language is a hypothesis which we select on the basis of evidence
21 combined with an inductive bias. Where are the replicators here? Who is doing the selection?
22 Give this latter perspective, the *neutral model* outlined in the previous section appears more
23 appropriate.

24 Which of these perspectives is correct? It is possible that in fact they are compatible – that they
25 are different ways of analysing the same process, namely social/cultural *adaptation*. The
26 challenge is in seeing how these analyses relate to one another and to the models that exist in the
27 literature.

28 Incidentally, we need to be clear that when we are talking about selection and replication here,
29 we are not talking about selection of heritable genetic variation (although that is clearly relevant
30 to language evolution, and to models of language evolution). Nor are we talking about the
31 *natural selection of cultural variants*, a mechanism by which fitter individuals are more likely to
32 survive and pass-on their cultural traits (although this too is likely to be important). Instead, we
33 are talking about the kind of adaptation that occurs purely through the complex process of
34 repeated cycle of utterance creation, interpretation, and internalisation that happens in language
35 transmission – whether it be in an iterated learning model focussing on vertical transmission, or a
36 negotiation model focussing on social coordination.

37 **Gene-cultural coevolution**

38 We pointed out in the Introduction that the current focus of language evolution modelling lies in
39 cultural evolution. It is however clear that a complete picture must integrate cultural with
40 biological (genetic) evolution. Formal modeling of gene-cultural coevolution began in the mid
41 1970s (Cavalli-Sforza & Feldman, 1981). Briscoe (2003, this volume) reviews models of gene-

1 culture co-evolution applied to language evolution. The basic idea is to use the population
2 geneticists' recursion equation formalism for cultural as well as genetic evolution. The result is a
3 system of linked dynamic equations that keep track of genes and culture as they change through
4 time. In general genes can influence culture via decision-making rules. An innate syntax might
5 constrain the evolution of languages. The flow of causation will in general also work the other
6 way. An element of a culturally transmitted protolanguage might exert selective pressure on the
7 genes. If genetic variation exists in the innate supports for language, and if more efficient
8 communication is favored, the variants that make the protolanguage more sophisticated will
9 increase in the population. Since cultural evolution will tend to be faster than genetic evolution,
10 cultural evolution will tend to be the driving partner in the coevolutionary circuit and genetic
11 evolution the rate limiting step. This does not tell us anything about the division of labor between
12 genes and culture at evolutionary equilibrium. That will depend upon many contingent costs and
13 benefits of transmitting adaptations genetically versus culturally. Very broadly speaking, the
14 genetic and cultural subsystems are both adaptive systems, and selection may be more or
15 indifferent as to how a given adaptation is transmitted.

16 While a complete coverage of gene-cultural modelling is a task for the future, one question
17 which has already been studied by gene-culture coevolutionists is whether and how the evolution
18 of various human adaptations may facilitate or constrain the evolution of language. Richerson
19 and Boyd (this volume) review models to explain human cooperation and symbolic boundary
20 marking. Language would seem to require a large measure of cooperation. Otherwise hearers
21 could not trust speakers. The non-cooperative case seems to exemplify the situation for most
22 other species. Hence communication systems in other animals are rather limited. Even in humans
23 people who live in different societies may not be trustworthy sources of information. Hence the
24 evolution of linguistic differences between human groups may be adapted to limit
25 communications from untrustworthy others.

26 **Advanced recurrent neural network models**

27 There exist a number of recurrent neural network architectures designed to model complex
28 linguistic (or visual) processing which are both computationally powerful and partly biologically
29 plausible. These models have not yet been used as a basis for evolution of language studies. Due
30 to their expressivity and the availability of advanced learning algorithms, they appear as
31 promising carrier formalisms for future evolutionary studies of grammatical processing.

32 The SHRUTI family of connectionist architectures (e.g., Shastri 1999) represents a long-standing
33 research strand to explain fast forward inferences in semantic text understanding. These models
34 are very complex, hand-designed networks of semantic and syntactic processing nodes which
35 communicate with each other by biologically motivated neural spike codes, enabling
36 combinatorial binding of different representation nodes across the network.

37 In machine learning, a recent landmark paper (Hinton & Salakhutdinov 2006) has unleashed a
38 flurry of research in *deep belief networks* (DBNs) and *restricted Boltzmann machines* (RBM).
39 With these models and novel learning algorithms, for the first time it has become feasible to train
40 deep conceptual hierarchical representations from large-volume real-life data in an unsupervised
41 way. While this field is so far preoccupied with visual learning tasks, the step toward speech/text
42 input is imminent (Y. Bengio, personal communication).

1 Even more recently, another family of hierarchical RNN-based models for learning
2 representations of very complex multiscale data is emerging. These models arise as
3 hierarchical/multiscale extensions of *Echo State Networks* (Jaeger & Haas 2004) or *Liquid State*
4 *Machines* (Maass, Natschläger & Markram 2002), the two main exemplars of a new
5 computational paradigm in neuroinformatics referred to as *Reservoir Computing* (Jaeger, Maass
6 & Principe 2005). They share a number of important characteristics with the neural models of
7 speech recognition explored by Peter F. Dominey (e.g. Dominey 2005; Dominey, Hoen & Inui
8 2006). In this field, language and speech modeling is indeed an important target domain.

9 An important characteristic of all recurrent neural network models, which sets them apart from
10 symbolic grammar formalisms, is that speech/language processing is construed as a fast, self-
11 organizing dynamical system, which does not need search subroutines and does not build interim
12 alternative interpretations. On the positive side, this leads to very fast processing (timescale of a
13 few neuronal delays); on the negative side, if an interpretation trajectory goes astray, this has to
14 be detected and separate repair mechanisms have to be invoked.

15 **Creatures-based modelling**

16 Simulation-based studies on language change today concentrate primarily on cultural
17 transmission dynamics. Neither brain structures nor genetic determinants for such brain
18 structures are modelled. This makes simulation-based research blind to some of the questions
19 that are raised in biological evolutionary accounts of the origins of language (Givón, this
20 volume; Fedor et al., this volume; Hilliard and White, this volume; Deacon, this volume). A
21 potentially very powerful avenue would be to simulate brain-body coevolution along the lines
22 staked out in *Artificial Life and Evolutionary Robotics* (e.g. Sims 1994, Nolfi & Floreano 2000,
23 Szathmary, 2007). In this research, artificial creatures are evolved in simulation or in physical
24 robotic hardware. A creature has a body equipped with sensors and actuators, and is controlled
25 by an artificial neural network that co-evolves with the body. Research of this kind has achieved
26 to evolve surprisingly complex and adaptive behaviour repertoires driven by neurocontrollers of
27 surprisingly small size. However, linguistic capabilities have so far largely fallen beyond the
28 scope of this research (but see Wischmann & Pasemann 2006). It appears a natural and
29 fascinating endeavour, albeit computationally challenging, to implement simulation scenarios
30 where body+brain creatures are evolved under selective pressures that favour efficient
31 communication. In this way, simulation-based research might tell an almost complete (if duly
32 simplified) story, connecting mechanisms of genetic coding of neural structures and the ensuing
33 slow "biological" adaptations with the fast cultural transmission dynamics that are the hallmark
34 of today's investigations.

35 **Detailed models of language learning during development**

36 Most models of language and syntax evolution treat an individual's learning of language during
37 their "childhood" in a very simplistic fashion. However, the transmission of language from one
38 generation to the next is clearly a central aspect of the evolution of language. Thus, more
39 elaborate modeling of the acquisition of language during infancy and childhood is called for.
40 Ideally, such models would take the embodied nature of language learning into account,
41 capturing how the learner interacts with their physical and social environment. At the same time
42 such models should be constrained by developmental psychology and developmental

1 neuroscience, providing constraints regarding the underlying neural structures, representations,
2 and learning mechanisms, as well as the nature of the language input that infants are typically
3 exposed to. So far, such approaches have been mostly restricted to learning early precursors of
4 language such as gaze following (Triesch et al., 2006) or learning of word meanings (Yu et al,
5 2005), but the time seems ripe to extend such models toward the acquisition of grammatical
6 structures.

7 **Case studies**

8 Scientific fields often organise their research around key challenges that are accepted by a large
9 group of researchers independently of the methods they are using. Also in technical fields, such
10 as machine learning, robotics, or high performance computing, there are often well-defined
11 challenges against which different research groups compete, often leading to very rapid progress
12 (as seen in the Robocup for example). What would such key challenges look like in the case of
13 language evolution? One possibility is to pick a certain domain which has been grammaticalised
14 in many languages of the world, although often in different ways, and show what cognitive
15 mechanisms and interaction patterns are needed to see the emergence of such a system in a
16 population of agents. Another possibility is to develop evolutionary models that are also capable
17 of capturing psycholinguistic data on actual human language behavior.

18 An example domain is the following: Many languages have ways to express predicate-argument
19 structure through a system of case grammar, either expressed morphologically or through word
20 order structure. The emergence of such a system requires not only the emergence of conventions
21 but also the emergence of the semantic and syntactic categories that underly it. A lot of data is
22 available from historical linguistics showing how such systems have arisen in human natural
23 languages, often by the grammaticalisation of verbs, and these data could constrain possible
24 models. There are already some attempts towards explaining the emergence of case grammar,
25 from the viewpoints of each of the paradigms introduced earlier (see Moy, 2006, Jäger 2007,
26 VanTrijp, 2007) and they can act as a starting point for tackling this challenge. It is not difficult
27 to imagine other aspects of grammar that could form the focus of well-defined challenges, and
28 once easier challenges are handled we could move to more challenging ones, such as clause
29 structure with long-distance dependencies.

31 **5 Conclusion**

32 Mathematical and computational models of language evolution make it possible to examine the
33 consequences of certain theoretical assumptions by mathematical deduction, large-scale
34 computer simulation, or robotic experimentation. Several efforts are under way to apply this
35 methodology to questions related to the problem of the origins and evolution of language. There
36 is not a single paradigm nor a single methodology, instead multiple paradigms explore different
37 questions. At this moment the models are mostly focusing on the origins of lexicons, categories
38 that can act as building blocks for conceptualisation, and simple languages with few of the
39 complex structuring principles found in human languages (but see Briscoe, this volume).
40 However we are confident that the technological foundations and mathematical tools will
41 become progressively more sophisticated and thus be able to tackle increasingly deeper and more
42 intricate question relating to the origins and evolution of syntax in language.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36

References

Aissen, J. (2003). Differential Object Marking: Iconicity vs. Economy. *Natural Language and Linguistic Theory* 21(3), 435-483.

Baronchelli, A., Dall'Asta, L., Barrat, A. and Loreto, V. (2006b). Topology induced coarsening in language games language. *Phys. Rev. E* 73, 015102(R)

Baronchelli, A., Felici, M., Caglioti, E., Loreto, V. and Steels, L. (2006a). Sharp transition towards shared vocabularies in multi-agent systems. *J. Stat. Mech.* P06014

Baronchelli, A., Loreto, V. and Steels, L. (2008). In-depth analysis of the Naming Game dynamics: the homogeneous mixing case. *Int. J. of Mod. Phys. C* 19, 785

Bergen, B. and Chang, N. (2004). Embodied Construction Grammar in Simulation-Based Language Understanding. In: J.-O. Östman and M. Fried (eds.), *Construction Grammar(s): Cognitive and Cross-Language Dimensions*. Johns Benjamins.

Boersma, P. (1998). *Functional Phonology: Formalizing the interactions between articulatory and perceptual drives*. Ph.D. dissertation University of Amsterdam.

Briscoe, E. J. (2000). *Grammatical Acquisition: Inductive Bias and Coevolution of Language and the Language Acquisition Device*. *Language* 76.2, 245-296.

Briscoe, E. J. (2003). Grammatical assimilation. In: Christiansen, M. and Kirby, S. (eds.), *Language Evolution: the States of the Art*. Oxford University Press

Cangelosi A., D. Parisi (1998). The emergence of a "language" in an evolving population of neural networks. *Connection Science*, 10(2), 83-9

Cavalli-Sforza, L. L. and Feldman, M. W. (1981) *Cultural Transmission and Evolution: A quantitative approach*. Princeton University Press.

Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT

Christiansen, M.H. (1992). The (non) necessity of recursion in natural language processing. In *Proceedings of the 14th Annual Cognitive Science Society Conference* (pp. 665-670). Hillsdale, NJ: Lawrence Erlbaum.

Christiansen, M.H. & Chater, N. (1999). Toward a connectionist model of recursion in human linguistic performance. *Cognitive Science*, 23, 157-205.

Christiansen, M.H. & Chater, N. (2003). Constituency and recursion in language. In M.A. Arbib (Ed.), *The handbook of brain theory and neural networks* (2nd ed., pp. 267-271). Cambridge, MA: MIT Press.

Christiansen, M.H. & Chater, N. (in press). Language as shaped by the brain. *Behavioral & Brain Sciences*.

Conway, C.M., & Christiansen, M.H. (2001). Sequential learning in non-human primates. *Trends in Cognitive Sciences*, 5, 539-546.

- 1 Dall'Asta, L., Baronchelli, A., Barrat, A. and Loreto, V. (2006a). Agreement dynamics on small-
2 world networks. *Europhys. Lett.* 73, 969
- 3 Dall'Asta, L., Baronchelli, A., Barrat, A. and Loreto, V. (2006b). Non-equilibrium dynamics of
4 language games on complex networks. *Phys. Rev. E* 74, 036105.
- 5 Croft, W. (2000): *Explaining Language Change: An Evolutionary Approach*. Longman, London.
- 6 De Beule, J. (2008). The Emergence of Compositionality, Hierarchy and Recursion in Peer-to-
7 Peer Interactions. In: Smith, A. D. M., Smith, K., and R. Ferrer i Cancho (2008). *The Evolution*
8 *of Language: Proceedings of the 7th International Conference (EVOLANG7)*. Singapore: World
9 Scientific Press.
- 10 Deutscher, G. 2005. *The Unfolding of Language: An Evolutionary Tour of Mankind's Greatest*
11 *Invention*. New York: Henry Holt
- 12 Dominey, P. F. (2005). From Sensorimotor Sequence to Grammatical Construction: Evidence
13 from Simulation and Neurophysiology. *Adaptive Behaviour* 13 (4), 347-361
- 14 Dominey, P. F., Hoen, M., Inui, T. (2006). A Neurolinguistic Model of Grammatical
15 Construction Processing. *Journal of Cognitive Neuroscience* 18 (12), 2088-2107
- 16 Dowman, M., Xu, J. and Griffiths, T.L. (2008). A Human Model of Colour Term Evolution. In
17 *The Evolution of Language*. Hackensack, NJ: World Scientific.
- 18 Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14, 179-211.
- 19 Galantucci, B. (2005) An Experimental Study of the Emergence of Human Communication
20 Systems. *Cognitive Science* (29) 737-767.
- 21 Gómez, R.L., & Gerken, L.A. (2000). Infant artificial language learning and language
22 acquisition. *Trends in Cognitive Sciences*, 4, 178-186.
- 23 Griffiths, T. L., Christian, B. R., & Kalish, M. L. (2008). Using category structures to test
24 iterated learning as a method for revealing inductive biases. *Cognitive Science*, 32, 68-107.
- 25 Griffiths, T. L., & Kalish, M. L. (2007). Language evolution by iterated learning with Bayesian
26 agents. *Cognitive Science*, 31, 441-480.
- 27 Haspelmath, M. (2007) Pre-established categories don't exist: consequences for language
28 description and typology. *Linguistic Typology* 11.1:119-132.
- 29 Heine, B. and Kuteva, T. (2008) *The Genesis of Grammar. A reconstruction*. Oxford University
30 Press, Oxford.
- 31 Hinton, G. E., Salakhutdinov, R. R. (2006). Reducing the Dimensionality of Data with Neural
32 Networks. *Science* 313 (July 28), 504-507
- 33 Hofbauer, J. & K. Sigmund (1998). *Evolutionary Games and Population Dynamics*. Cambridge
34 University Press, Cambridge (UK).
- 35 Jäger, G. (2004). Learning Constraining Subhierarchies: The Bidirectional Gradual Learning
36 Algorithm. In R. Blutner & H. Zeevat (eds.), *Optimality Theory and Pragmatics*. Palgrave
37 MacMillan, 251-287.

- 1 Jäger, G. (2007). Evolutionary Game Theory and Typology: A Case Study. *Language* 83(1), 74-
2 109.
- 3 Jäger, G. (2008a). Evolutionary stability conditions for signaling games with costly signals.
4 *Journal of Theoretical Biology* 253(1), 131-141.
- 5 Jäger, G. (2008b). The evolution of convex categories. *Linguistics and Philosophy* 30(5), 551-
6 564.
- 7 Jäger, G. & R. van Rooij (2007). Language structure: psychological and social constraints.
8 *Synthese* 159(1), 99-130.
- 9 Jaeger, H., Haas, H. (2004). Harnessing Nonlinearity: Predicting Chaotic Systems and Saving
10 Energy in Wireless Communication. *Science* 304, 78-80
- 11 Jaeger, H., Maass, W., Principe, J. (2007). Special issue on echo state networks and liquid state
12 machines. *Neural Networks* 20 (3), 287-289
- 13 Kalish, M. L., Griffiths, T. L., & Lewandowsky, S. (2007). Iterated learning: Intergenerational
14 knowledge transmission reveals inductive biases. *Psychonomic Bulletin and Review*, 14, 288-
15 294.
- 16 Kaplan, R. and J. Bresnan, (1982) LFG: a Formal System for Grammatical Representation, in
17 Bresnan, ed. *The Mental Representation of Grammatical Relations*, MIT Press, Cambridge,
18 1982, 173-281.
- 19 Kimura, M. (1983). *The Neutral Theory of Molecular Evolution*. Cambridge University Press,
20 Cambridge.
- 21 Kirby, S. (2006). Major Transitions in the Evolution of Language. In Angelo Cangelosi, Andrew
22 D. M. Smith and Kenny Smith (eds.), *The Evolution of Language*. World Scientific.
- 23 Kirby, S., Dowman, M., & Griffiths, T. (2007). Innateness and culture in the evolution of
24 language. *Proceedings of the National Academy of Sciences*, 104, 5241-5245.
- 25 Komarova, N.L., Niyogi, P. and Nowak, M.A. (2001). Evolutionary dynamics of grammar
26 acquisition. *Jour. Theor. Biol.*, 209(1), pp. 43-59.
- 27 Komarova, N.L. and Nowak, M.A. (2001). Population dynamics of grammar acquisition, in A.
28 Cangelosi & D. Parisi (eds.) *Simulating the Evolution of Language*, pp. 149-164, Springer
29 Verlag, London.
- 30 Komarova, N.L. and Nowak, M.A. (2003). Language, learning and evolution, in S. Kirby & M.
31 Christiansen (eds.) *Language Evolution: The States of the Art*, pp. 317-337, Oxford University
32 Press.
- 33 Komarova, N.L. and Jameson, K.A. (2008). Population heterogeneity and color stimulus
34 heterogeneity in agent-based color categorization. *Jour. Theor. Biol.* 253(4), pp. 680-700.
- 35 Jameson, K.A. and Komarova, N.L. (2009). Evolutionary models of color categorization:
36 investigations based on realistic population heterogeneity. Submitted to *JOSA*.
- 37 Labov, W. (2001). *Principles of Linguistic Change: Social Factors*. Edited by P. Trudgill. Vol.
38 29, *Language in Society*. Malden, MA: Blackwell

- 1 Maass, W., Natschläger, T., Markram, H. (2002). Real-time Computing Without Stable States: A
2 New Framework for Neural Computation Based on Perturbations. *Neural Computation* 14 (11),
3 2531-2560
- 4 MacDonald, M.C. & Christiansen, M.H. (2002). Reassessing working memory: A comment on
5 Just & Carpenter (1992) and Waters & Caplan (1996). *Psychological Review*, 109, 35-54.
- 6 Maynard Smith, J. (1982). *Evolution and the Theory of Games*. Cambridge University Press,
7 Cambridge (UK).
- 8 McElreath, R., Boyd, R. and Richerson, P. J. (2003). Shared norms and the evolution of ethnic
9 markers. *Current Anthropology* 44:122-129.
- 10 McElreath, R., & Henrich, J. (2008). Modeling Cultural Evolution. In R. Dunbar & L. Barrett
11 (Eds.), *Oxford Handbook of Evolutionary Psychology* (pp. 571-585). Oxford Univ Press.
- 12 Mesoudi, A. (2007). Using the methods of experimental social psychology to study cultural
13 evolution. *Journal of Social, Evolutionary and Cultural Psychology* 1(2), 35-58.
- 14 Mezard, M, Parisi, G. and Virasoro, M.A. (1987). *Spin Glass Theory and Beyond*. World
15 Scientific (Teaneck, NJ-USA).
- 16 Moy, J. (2006). *Word Order and Case in Models of Simulated Language Evolution*.
17 Ph.D. thesis, University of York.
- 18 Nolfi, S., Floreano, D. (2000). *Evolutionary Robotics: The Biology, Intelligence, and*
19 *Technology of Self-Organizing Machines*. MIT Press
- 20 Nowak, M.A., Komarova, N.L. and Niyogi, P. (2001). Evolution of universal grammar. *Science*,
21 291, pp. 114-118.
- 22 Nowak, M.A., Komarova, N.L. (2001). Toward an evolutionary theory of language. *Trends in*
23 *Cognitive Sciences*, 5(7), pp. 288-295.
- 24 Nowak, M. A. and Krakauer, D. C. (1999). The evolution of language. *Proceedings of the*
25 *National Academy of Sciences*, 96(14):8028–8033.
- 26 Nowak, M. A., Krakauer, D. C. and Dress, A. (1999). An error limit for the evolution of
27 language. *Proceedings of the Royal Society: Biological Sciences* 266, 2131–6.
- 28 Pickering, M. and S. Garrod, S. (2004). Toward a Mechanistic Psychology of Dialogue.
29 *Behavioral and Brain Sciences*. 27(2):169–226
- 30 Puglisi, A., Baronchelli, A., and Loreto, V. (2008). Cultural route to the emergence of linguistic
31 categories, *Proc. Natl. Acad. Sci. USA* 105, 7936
- 32 Reali, F. & Christiansen, M.H. (in press). Sequential learning and the interaction between
33 biological and linguistic adaptation in language evolution. *Interaction Studies*.
- 34 Reali, F., Christiansen, M.H. & Monaghan, P. (2003). Phonological and distributional cues in
35 syntax acquisition: Scaling up the connectionist approach to multiple-cue integration. In
36 *Proceedings of the 25th Annual Conference of the Cognitive Science Society* (pp. 970-975).
37 Mahwah, NJ: Lawrence Erlbaum.
- 38 Sag, I., T. Wasow, and E. Bender (2006). *Syntactic Theory. A Formal Introduction*. University

- 1 of Chicago Press, Chicago.
- 2 Shastri, L. (1999). Advances in Shruti – a neurally motivated model of relational knowledge
3 representation and rapid inference using temporal synchrony. *Artificial Intelligence* 11, 79-108
- 4 Sims, K. (1994). Evolving Virtual Creatures. *Proc. of Siggraph 1994*, 15-22
- 5 Steels, L. (1995). A Self-organizing Spatial Vocabulary. *Artificial Life* 2(3), pp. 319-332.
- 6 Steels, L. (2007). The Recruitment Theory of Language Origins. In: Lyon, C. and Nehaniv, C. L.
7 and Cangelosi, A. (ed.) *Emergence of Language and Communication*, pages 129-151, Springer.
8 Berlin, 2007.
- 9 Steels, L. and T. Belpaeme (2005). Coordinating Perceptually Grounded Categories through
10 Language. *A Case Study for Colour. Behavioral and Brain Sciences*. 24(6).
- 11 Steels, L. and Beule, J. D. (2006). Unify and Merge in Fluid Construction Grammar. In P. Vogt
12 and et al., editors, *Symbol Grounding and Beyond: Proceedings of the Third International*
13 *Workshop on the Emergence and Evolution of Linguistic Communication*, pages 197-223.
14 Springer.
- 15 Steels, L. and M. Loetzsch (2008). Perspective Alignment in Spatial Language. In: Coventry, K.,
16 J. Bateman and T. Tenbrink (eds) *Spatial Language in Dialogue*. Oxford University Press.
17 Oxford.
- 18 Szathmáry, E. (2000). The evolution of replicators. *Philosophical transactions of the Royal*
19 *Society of London. Series B, Biological sciences*, Vol. 355, No. 1403. pp. 1669-1676.
- 20 Szathmáry, E. *et al.* (2007). *In silico* evolutionary developmental neurobiology and the origin of
21 natural language. In: (C. Lyon, C. Nehaniv and A. Cangelosi, eds) *Emergence of communication*
22 *and language*. Springer, London, pp. 151-187.
- 23 Trapa, P. E., Nowak, M. A. (2000). Nash equilibria for an evolutionary language game. *Journal*
24 *of Mathematical Biology* 41.172–88.
- 25 Triesch, J., Teuscher, C., Deak, G, and Carlson, E. (2006). Gaze Following: why (not) learn it?
26 *Developmental Science*, 9(2):125-147
- 27 van Rooij (2004a). Evolution of conventional meaning and conversational principles. *Synthese*
28 139, 331-366.
- 29 van Rooij (2004b). Signalling games select Horn strategies. *Linguistics and Philosophy* 27, 493-
30 527.
- 31 van Rooij, R. (2008). Evolutionary motivations for semantic universals. In R. Eckardt, G. Jäger
32 & T. Veenstra, eds., *Variation, Selection, Development: Probing the Evolutionary Model of*
33 *Language Change*, Mouton de Gruyter.
- 34 Van Trijp, Remi (2008). The Emergence of Semantic Roles in Fluid Construction Grammar. In:
35 Smith, Andrew D.M. and Kenny Smith and Ramon Ferrer i Cancho (eds.) *The Evolution of*
36 *Language. Proceedings of the 7th International Conference*. p. 346-353. Singapore: World
37 Scientific Publishing.

- 1 Wörneryd, K. (1993). Cheap talk, coordination and evolutionary stability. *Games and Economic*
2 *Behavior* 5, 532–46.
- 3 Wellens, P., Loetzsch, M. and Steels, L. (2008). Flexible Word Meaning in Embodied Agents.
4 *Connection Science*, 20 (2), 173-191.
- 5 Wischmann, S., Pasemann, F. (2006). The Emergence of Communication by Evolving
6 Dynamical Systems. In: *From animals to animats 9: Proceedings of the Ninth International*
7 *Conference on Simulation of Adaptive Behaviour*, edited by Nolfi, S. and Baldassarre, G. and
8 Calabretta, R. and Hallam, J. and Marocco, D. and Meyer, J.-A. and Parisi, D.. Springer Verlag,
9 777-788
- 10 Yu, C., Ballard, D., Aslin, R. (2005). The Role of Embodied Intention in Early Lexical
11 Acquisition. *Cognitive Science* 29(6):961-1005
- 12 Zuidema, W. (2002). Language adaptation helps language acquisition - a computational model
13 study. In: Hallam B., Floreano D., Hallam J., Hayes G., and Meyer J-A. (eds.), *From Animals to*
14 *Animats 7: Proceedings of the Seventh International Conference on Simulation of Adaptive*
15 *Behavior*.