Modeling Interactions Between Language Evolution and Demography

PAUL VOGT[1]

Abstract In this article I provide a review of studies that have modeled interactions between language evolution and demographic processes. The models are classified in terms of three different approaches: analytical modeling, agent-based analytical modeling, and agent-based cognitive modeling. I show that these approaches differ in the complexity of interactions that they can handle and that the agent-based cognitive models allow for the most detailed and realistic simulations. Thus readers are provided with a guideline for selecting which approach to use for a given problem. The analytical models are useful for studying interactions between demography and language evolution in terms of high-level processes; the agent-based analytical models are good for studying such interactions in terms of social dynamics without bothering too much about the cognitive mechanisms of language processing; and the agent-based cognitive models are best suited for the study of the interactions between the complex sociocognitive mechanisms underlying language evolution.

The question of how human languages originated, spread over the globe, and are constantly changing remains one of the biggest challenges for 21st-century scientists (e.g., Christiansen and Kirby 2003). In the past few decades these challenges have attracted increasing attention from the scientific community. Among the major driving forces behind this increased attention has been improved computational capacity and computerized techniques. In this paper I am concerned with applying demographic models and data to computer simulations of language evolution.

One of the reasons for the success of computer modeling in the field of language evolution is that theoretical models, whose results are often complex and consequently hard to predict on the back of an envelope, can be simulated and their results can be compared with empirical observations, allowing one to validate the theory. Here also lies a weak point of the approach: Many models are highly abstract and simplified such that a comparison with empirical data (e.g., linguistic or psychological) is often difficult to achieve (see, e.g., Vogt and de Boer 2010).

[1]Tilburg Center for Creative Computing, Tilburg University, P.O. Box 90153, Tilburg, The Netherlands.

Human Biology, April–June 2009, v. 81, nos. 2–3, pp. 237–258.

KEY WORDS: LANGUAGE EVOLUTION, LANGUAGE ORIGINS, LANGUAGE CHANGE, LANGUAGE CONTACT, ANALYTICAL MODELS, AGENT-BASED MODELS.

Although some comparisons of outcomes of computer models with empirical data have been successful (e.g., de Boer 2001; Griffiths et al. 2008; Nettle 1999c), many computational studies still need proper grounding in empirical studies.

Surprisingly, archaeology is a discipline that has received little attention in the computational modeling of language evolution, although increasingly more empirical data have become available, in particular regarding demographic dynamics (Klein 2000; Mellars 2005). Most theoretical models that take demographic data into account are analytical models (e.g., Nowak et al. 2001). Although these models have helped to gain many insights into the evolution of language, they have their limitations. An alternative to analytical modeling is agent-based modeling, in which a population of individuals interact with each other and develop aspects of language by learning from each other [for a review of analytical and agent-based models of language *competition,* see Kandler (2009) (this issue)]. However, so far few agent-based models (the type of models focused on in this paper) take demographic data into account, but see, for example, Briscoe (2000), Nettle (1999a), and Parisi et al. (2008) for counterexamples.

In this paper I discuss some of the limitations of analytical modeling and put forward the arguments in favor of an agent-based modeling approach to simulate the relations between demography and language evolution. The purpose of this paper is not to dismiss analytical modeling as a tool but rather to promote agent-based modeling as an additional (and potentially more powerful) tool. I introduce two classes of agent-based models: agent-based analytical models (e.g., Baxter et al. 2009; Nettle 1999a), in which individuals are modeled using mathematical equations; and agent-based cognitive models (e.g., Briscoe 2000; Parisi et al. 2008; Vogt, 2007a), in which individuals are implemented with a cognitive model. My objective is to provide readers who are interested in studying the interactions between demography and language evolution with a guideline for selecting one of these modeling approaches on the basis of their characteristics, advantages, and disadvantages.

In the next section I review a number of studies that have used demographic models and/or data in analytical models, agent-based analytical models (ABAMs), or agent-based cognitive models (ABCMs). After that, I present a few case studies of ABCMs that have used empirical (demographic) data and those that have not. The discussion focuses on the advantages and disadvantages of ABCMs over analytical models and ABAMs and outlines some ideas for future research. I argue that ABCMs are most realistic, but at the expense of increased (computational) complexity.

Models of Language Evolution

One of the aims of studying models of language evolution is to understand the effect of demographic processes on issues such as language change, language death, and language creation. In this section I review three methods for modeling language evolution, going from macroevolutionary analytical or analytical models to microevolutionary agent-based models, which are further subdivided into agent-based analytical models and agent-based cognitive models.

Analytical Models. In line with standard models of cultural evolution (Boyd and Richerson 1985; Cavalli-Sforza and Feldman 1981), various analytical models have been proposed to investigate how certain linguistic phenomena, such as whole languages, can propagate in a population and under what conditions stable (or unstable) linguistic systems can emerge. Such models involve, for instance, dynamic system models (Abrams and Strogatz 2003; Niyogi and Berwick 1996) or Markov chains of Bayesian learners (Kirby et al. 2007).

A typical analytical model of language evolution contains a differential equation that quantifies the dynamics of the frequency with which a certain linguistic aspect (e.g., syntax, grammar, or phonemic element) occurs in some population of speakers (e.g., Komarova et al. 2001; Lightfoot 1999; Niyogi and Berwick 1996; Nowak et al. 2001). An equation that is often used is the differential equation of Nowak and colleagues (Nowak et al. 2001; Komarova et al. 2001):

$$\frac{\partial x_i}{\partial t} = \sum_{j=1}^{n} x_j f_j Q_{ji} - \phi x_i, \qquad i = 1, \ldots, n, \tag{1}$$

where $\phi = \sum_i x_i f_i$ is the average fitness or grammatical coherence of the population, n is the number of different variants (e.g., grammars), x_i is the frequency of variant i (i.e., the number of users of variant i), f_i is the fitness of a variant, and Q_{ji} is a transmission matrix that specifies how likely a certain variant will be transmitted from one generation to the next correctly (Q_{ii}) or incorrectly (Q_{ji}, $i \neq j$).

Another popular model assumes two fixed languages that compete with each other (Abrams and Strogatz 2003; Kandler and Steele 2008; Minett and Wang 2008; Patriarca and Leppänen 2004). This model calculates the dynamics of the number of speakers of a certain variant. An example of such a model is the reaction-diffusion system of Kandler and Steele (2008), who formulated a competition between two languages by extending the earlier models of Abrams and Strogatz (2003) and Patriarca and Leppänen (2004):

$$\frac{\partial u_A}{\partial t} = d_A \Delta u_A + r_A u_A \left(1 - \frac{u_A}{K - u_B}\right) + c u_A u_B, \tag{2a}$$

$$\frac{\partial u_B}{\partial t} = d_B \Delta u_B + r_B u_B \left(1 - \frac{u_B}{K - u_A}\right) + c u_A u_B. \tag{2b}$$

This reaction-diffusion system describes the behavior of two variables u_A and u_B, which give the frequencies of users of language A or B at a given time t. The behavior is governed by six parameters—d_A, d_B, r_A, r_B, K, and c, which relate to a spatial dispersal of the languages (d_i), intrinsic growth of a language population (r_i, K), and the social status differences of both languages ($c > 0$). The model's

Table 1. Number of Variables and Other Parameters Used in Various Analytical Models

Study	*Number of Variables*	*Number of Other Parameters*
Abrams and Strogatz (2003)	2	4
Kandler and Steele (2008)	2	6
Minett and Wang (2008) (analytical model)	2 (3)	9
Nowak et al. (2001)	n	$n \times n$ (2)
Patriarca and Leppänen (2004)	2 (4)	6

analysis shows that coexistence of both languages is impossible but that the time it takes that the lower status language (B) to die out is crucially dependent on the six parameter settings (Kandler and Steele 2008). For a more extensive review of analytical models of language competition that also considers agent-based approaches, see Kandler (2009) (this issue).

The advantage of such analytical models is that they tend to be simple; that is, only a few variables and parameters influence the system (see Table 1). When the system behaves similarly to empirically observed systems (e.g., Abrams and Strogatz 2003; Niyogi and Berwick 1996), the model provides a possible explanation in terms of the parameters that define the model. Given the complexity of the real systems (languages, language populations, and individual speakers), the model can describe only macroevolutionary effects. Although this is an asset, it also limits the model in its explanatory power. For instance, analytical models are deterministic such that they behave in a unique way given the initial conditions. It is arguable whether the deterministic nature of these models is sufficiently capable to describe and capture the complex dynamics of language evolution [see, e.g., Briscoe (2000) for a neat discussion]. In the words of Minett and Wang (2008): "[These models do] not trace the states of every single speaker, only the proportions of speakers having each state. Thus, it specifies a model of the *expected* behaviour of the competition, but not the *range* of behaviours that can result from a given initial state or their relative likelihoods" (p. 34, emphasis in the original).

Table 1 shows the number of variables and parameters used in some of the analytical models of language evolution discussed in this paper. Note that the data for Minett and Wang (2008) refer to their analytical model; they have three variables, but they conveniently have set one of them to 0, thus using only two variables. The Nowak et al. (2001) model has n variables, one for each possible variant. Although theoretically they assume $n \times n$ parameters, these are defined in only two parameters. Patriarca and Leppänen (2004) have used four variables in their model, but this can be reduced to two because of a direct dependency between them.

Agent-Based Analytical Models. The alternative that Minett and Wang (2008) propose is to model each individual by letting him or her learn a given variant based, for instance, on the likelihood that the variant will be learned. The model they propose is what can be classified as an agent-based analytical model, in which each individual agent is characterized by a mathematical formula. ABAMs have also been adopted by various other researchers (e.g., Baxter et al. 2009; de Boer 2005; Nettle 1999c).

Baxter et al. (2009), for instance, modeled the evolution of two linguistic variants α and β by defining how interactions between two individuals i and j (selected according to some network structure G_{ij} bearing geographic and social properties) change the likelihood that these agents will use a given variant (see also Baxter et al. 2006):

$$x_i(t+1) = \frac{1}{C}\left[x_i(t) + \lambda\frac{n}{T} + \lambda H_{ij}\frac{m}{T}\right], \tag{3}$$

where $x_i(t)$ is the likelihood that agent i uses one variant, say α, at time t. The parameter λ is a weight that "can be thought of as the receptiveness of the speaker to changing their grammar on the basis of the language they hear" (Baxter et al. 2009). The agents interchange in a given round T (10 or 20) tokens: $a \leq T$ tokens of one variant, α, and $T - a$ tokens of the other variant, β. Of the a tokens, n are produced by agent i and m are produced by agent j. The likelihood of agent i adopting the token produced by agent j is based on the social network status H_{ij} between both individuals. The factor $C = 1 + \lambda + \lambda H_{ij}$ is used to normalize the equation. Agent j also adopts its likelihood of using variant α by the same equation, but with the indexes i and j as well as the numbers n and m interchanged. See also Baxter et al. (2006) for more details of the model.

This system was run for a large number of interactions between different individuals, and, after initializing their system using demographic information on the immigration and population dynamics of New Zealand, Baxter et al. (2009) showed how variants of New Zealand English evolved in their model similarly to the way it has evolved in reality according to the theory of Trudgill (2004). In particular, the results of their model suggest that the social network information stored in H_{ij} does not influence the spread of a variant.

It should be clear that the model of Baxter et al. (2009) is far more detailed in defining aspects of learning than the analytical models discussed earlier. This is not because the number of parameters is much larger but because the variants are calculated for each individual interaction, increasing the number of variables in the systems proportionally to the population size N. This is a general aspect of ABAMs (see Table 2), but also of ABCMs, which are discussed next.

Table 2 shows the number of variables and parameters used in some of the ABAMs discussed in this paper. In these models the number of variables tends to be proportional to the population size N. Theoretically, the number of parameters can also be proportional to N (Nettle 1999c) or even $N \times N$ (Baxter et al. 2009) to define aspects such as the likelihood that different agents will interact with each other (often these parameters tend to have the same values for practical reasons).

Agent-Based Cognitive Models. Even more detailed simulations on language evolution can be obtained through agent-based cognitive models, in which the

Table 2. Number of Variables and Parameters Used in Various Agent-Based Analytical Models

Study	*Number of Variables*	*Number of Parameters*
Nettle (1999c)	$\sim N$	$2N + 4$
Baxter et al. (2009)	$3N$	$N \times N + 3$
Minett and Wang (2008) (ABAM)	$2N$	9

agents are not specified by a mathematical formula but by a computational cognitive model. These cognitive models typically contain mechanisms for processing language production, language interpretation, language learning, and memory. Some of these models are implemented in robots that have an adaptive categorization module with which they categorize the real world or a virtual one [for an overview see, e.g., Vogt (2006)], but most models are implemented in computer simulations with predefined semantics (Cangelosi and Parisi 2002; Kirby 2002). The linguistic properties that these agents learn vary from phonetic systems, such as vowels (de Boer 2001; Livingstone and Fyfe 2000) or syllables (Oudeyer 2005), to lexicons and vocabularies (Baronchelli et al. 2006; Kaplan 2005; Oliphant 1996; Steels 1996; Vogt and Coumans, 2003), syntax (Batali 1998; Briscoe, 2000; Kirby et al. 2004), or grammar (Steels 2005; Vogt 2005a).

Computational studies of the evolution of language often use the language game model introduced by Steels (1996) or a variant of it. In such a model there is a population of N individuals (or agents), whose aim is to develop a shared communication system by engaging in language games. Typically, the population starts from scratch, meaning that none of the agents start with any language. A language game is played by two randomly (or otherwise) selected agents: a speaker and a hearer. Note that it is possible to control the likelihood that two agents will interact by keeping track of the social networks of the agents (see, e.g., Dall'Asta et al. 2006; Gong et al. 2008; Ke et al. 2008). The agents are usually set in a context that contains some objects and/or events represented by meanings that agents develop ontogenetically or that the designer predefined. The speaker selects one object or event as the topic and produces an expression (a set of signals or words) to convey that topic. In turn, the hearer tries to interpret the expression, ideally identifying the intended topic. At the end (or sometimes during) a language game, the agents adapt their private linguistic competences (lexicon and/or grammar) and are thus learning from the interaction.

Early during development, an agent is likely to lack a way of expressing a meaning, because the agent starts without knowing any words. When a speaker consequently fails to produce an expression, it invents a new expression (typically an arbitrary string of letters from a finite alphabet). Likewise, when the hearer does not know a word or fails to identify the intended meaning (which may be conveyed through extralinguistic communication, e.g., using pointing gestures), the hearer can add the received expression to its memory and associate it with the

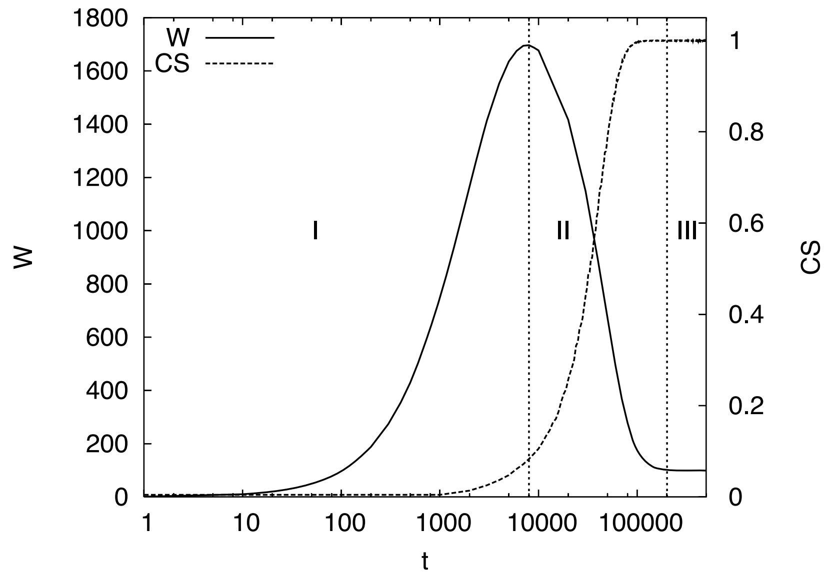

Figure 1. The three phases of language development for the language game model with a population of 50 agents and 100 meanings. In phase I the population invents many words (*W*), leading to many synonyms in the language. During this phase, communicative success (*CS*) remains low. In phase II the synonyms are dampened by the positive feedback loop with which the weights are adapted during the language games, leading to a sharp transition toward communicative successful lexicons. In phase III a stable lexicon has been established, in which the number of words equals the number of meanings, that is, $W = M$, and *CS* has converged to 1.

intended meaning. Agents can acquire multiple (i.e., many-to-many) mappings between signals and meanings that then compete with each other for being used. This competition is usually facilitated by a weight, which is increased when an association is used successfully in a language game or decreased when the association is used wrongly or when it is competing with an association that was used successfully. Lowering the weights of competing associations is called lateral inhibition, which is common to self-organizing maps and Hebbian learning. When agents have to choose between different associations to express a meaning or interpret a word, they will use the association that has the highest weight.

The population of agents play a relatively large number of such language games, yielding three phases in the evolution (see Figure 1): (I) a rise of synonymy, (II) a damping of synonymy, and (III) a stable communication system. During phase I of the simulation, speakers invent a new word for a meaning whenever they wish to produce an expression to convey this meaning even though they had not yet acquired a word for it. Because in these models agents play language games in pairs of randomly selected agents about arbitrarily selected meanings, it takes a while before one word is spread across the entire population; in fact, it has been shown analytically that spreading takes about $N \log N$ games (Kaplan 2005). Early during the evolution, there is a large chance that an agent wants to convey a meaning it has no word for, so it invents a new word. After some time, each individual has acquired a word for each meaning, either through invention or adoption. At this time, a maximum number of words are used in the global language and phase I ends.

When the maximum number of words in the population is reached, phase II starts, in which the competition dynamics of association weights dampens the synonymy that arose in the global language. Agents hear and adopt new words, which then compete through the update of association weights. The continuing strengthening of successfully used associations and the lateral inhibition of competing ones mean that, at the global level, more agents tend to use the same words for expressing the same meanings, thus causing a positive feedback loop and improving communicative success.

The process during phase II is much related to Darwinian selection at the cultural level, because those elements that are used successfully tend to be reused more often and thus have an increasing chance of becoming successful, which in turn strengthens those elements, causing a positive feedback loop that leads to the emergence of a stable attractor. As in many evolutionary systems, variation is crucial to facilitate, for instance, the emergence of compositionality in language (Vogt 2007b), as I discuss in the next section.

Once all agents use the same words to express the same meanings, communicative success is 100% and phase III starts. In this final phase the system is stable, because the same word-meaning mappings are used by all agents successfully, so they are always strengthened. Only when new agents or meanings enter the system can the global lexicon change. This can also happen for some other reason, for instance, when noise appears in the transmission. The reason for such changes is that the introduction of new agents or meanings tends to induce new variation in the language, which selection can act on.

Baronchelli et al. (2006) discovered that the time it takes for the maximum number of words to have emerged ($t_{\max}$) is proportional to the time of convergence T_c (i.e., the time it takes for phase III to commence). Moreover, they found that the maximum number of words ($W_{\max}$) acquired by the population at time $t_{\max}$ is also proportional to the time of convergence. So

$$T_c \propto t_{\max} \propto W_{\max}. \tag{4}$$

Kaplan (2005) and Baronchelli et al. (2006) have found that the time of convergence T_c depends on the group size N:

$$T_c \propto N^{\beta}, \tag{5}$$

where β is a constant. Given a value of $\beta \approx 1.3$ and Kaplan's analytical derivation of the time it takes for words to spread in the population, Eq. (5) suggests that

Table 3. Number of Variables and Parameters Used in Various Agent-Based Cognitive Models

Study	*Number of Variables*	*Number of Parameters*
Kaplan (2005)	$N \times N \times W$	C
Baronchelli et al. (2006)	$N \times W$	C
Vogt (2007a)	$N \times g \times M' \times W'$	C
Briscoe (2000)	$N \times (g + 4)$	C
Parisi et al. (2008)	$2N$	$6N$

there is, indeed, an $N \log N$ dependency for T_c. Note that Baronchelli et al. (2006), using a slightly different model, found a different exponent, namely, $\beta \approx 1.5$. The difference was that Baronchelli et al.'s model assumed only one meaning and allowed only one-to-one mappings between the meaning and signal, whereas Kaplan's model allowed a competition between words. Apparently, competition between words is beneficial for the convergence of the system.

Time of convergence in the language game is dependent not only on group size but also on the number of meanings M to be named. Kaplan (2000) discovered a linear dependency, such that

$$T_c \propto MN^{\beta}. \tag{6}$$

What is clear from this example is that the complexity of the model in terms of the number of variables substantially exceeds the complexity of the (agent-based) analytical models (Table 3). In this example each agent constructed a matrix specifying the number of meanings M and the (variable) number of words W an agent acquired individually. Each cell in this matrix contains a variable indicating the weight of the association. So, each agent i stores $M \times W_i$ variables, yielding a total of $\sum_i MW_i$ variables. Because the end result yields a vocabulary in which each agent uses the same distinct word for each distinct meaning, the total number of variables V in this system is $V \geq N \times M \times M$.

The complexity of an ABCM becomes even more apparent when one realizes the added complexity in the number of mechanisms that need to be implemented in order to process the production, interpretation, and learning of language. These processes might involve relatively complicated and nonlinear decision processes, thus increasing the complexity with respect to the analytical models and the ABAMs. Because it is hard to quantify this complexity, it is not included in the analysis presented in Table 3.

Table 3 shows the number of variables and parameters used in some ABCMs. In these models the number of variables relates to the population size N, the number of meanings M, the number of words W, and/or the grammar size g. The language game model introduced by Steels (1996), the example from Figure 1, and the study of Steels and McIntyre (1999), to be discussed in the next section, all have about the same complexity as Kaplan's (2005) model. The M' and W' regarding Vogt's (2007a) model refer to the number of meanings and words used for each grammar rule. Note that in the Parisi et al. (2008) study, *N* does not refer to group size but is proportional to the size of Europe, as discussed later. The number of parameters in these examples is usually a small constant *C* but may also depend on *N*, as in Parisi et al.'s (2008) study.

Some Case Studies of ABCMs

It is possible to distinguish between studies that do not use empirical data on demography and those that do use empirical data. In their ABCMs, Baronchelli et al. (2006) and Kaplan (2005) looked at demographic effects but did not use any empirical data. Rather, they investigated, hypothetically, what would happen if populations had different sizes. Baxter et al.'s (2009) ABAM did use empirical demographic data to initialize the model, and the investigators compared the outcome to empirical linguistic data. In this section I examine a few case studies of ABCMs in which demography has an effect on language evolution. Attentive readers might miss a detailed discussion of the renowned work by Kirby and colleagues (e.g., Kirby 2001; Smith et al. 2003). The reason for this is that, although Kirby and co-workers have made extensive use of ABCMs, their studies have typically involved a population of only two agents. Hence they have hardly studied any demographic effects on language evolution.

Demographic Effects Without Empirical Data. In line with the Baronchelli et al. (2006) and Kaplan (2005) studies, Vogt (2007a) studied the effect of different population sizes on the emergence of compositionality in language. Compositionality refers to expressions in which parts of an expression relate to parts of its meaning and the way these are combined. The agents were initialized with a combinatorial meaning space (e.g., specifying objects by their colors and shapes) and played language games in the same manner as before. In this study, however, agents were given a mechanism to discover and use regular patterns in both signals and meanings that allowed them to construct compositional structures such as *red circle*, *red square*, *blue circle*, and so on. In this model the agents initially developed a holistic communication system, in which each combinatorial meaning was associated with a distinct random signal (i.e., no part of the expression relates a part of its meaning). Gradually, when new generations were introduced and older ones were removed, the language became more and more compositional.

Vogt (2007a) found that for larger population sizes the amount of compositional structures that emerged in the languages increased, up to a certain limit (see Figure 2a). Like Baronchelli et al. (2006) and Kaplan (2005), Vogt (2007a) discovered a relation between group size and time of convergence, which essentially had three phases (see Figure 2b). In the first and third phases relations were a

b

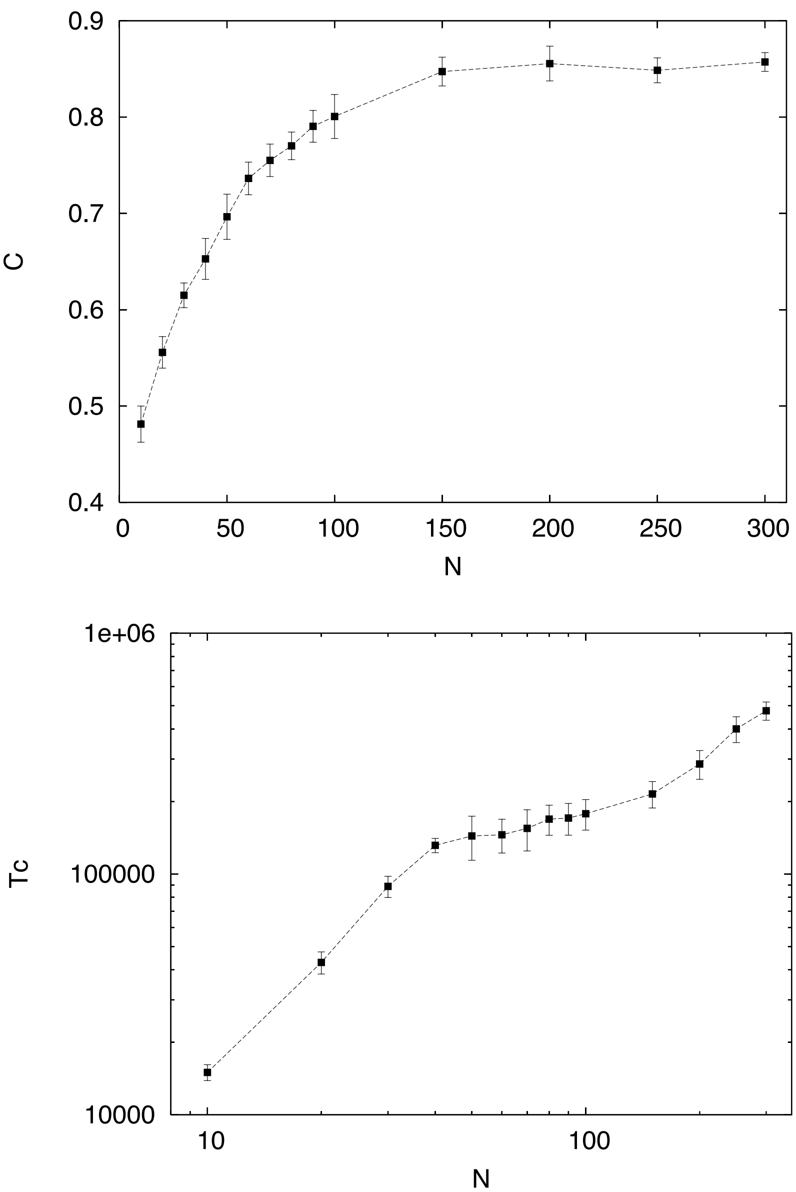

Figure 2. (a) Proportion of compositional expressions C used in a language when they develop for different group sizes N. After saturation, about 85% of the expressions produced or interpreted had a compositional representation, whereas 15% were holistic. (b) Time of convergence T_c for different group sizes N. Reprinted from Vogt (2007a) with kind permission of Springer Science + Business Media.

similar to those found for the emergence of vocabularies (i.e., $T_c \propto MN^{\beta}$), but in between there was a transition. This transition can be explained by the observation that, when the language becomes more compositional, fewer meanings can be spread among the population, thus reducing T_c.

The reason that compositionality emerged more substantially among larger group sizes can be explained by observing that in larger populations more words entered the linguistic community [cf. Eqs. (4) and (6)]. Because the number of possible letters with which words were created was limited, the probability that an individual would find regular patterns in both words and meanings increased as well. The discovery of such regular patterns drove the emergence of compositionality, which was further reinforced by the positive feedback loop inherent to the language game. In this way, larger groups caused an increased variation in the linguistic elements forming the language, which increased the chance of finding good compositional traits that would lead to the emergence of a smaller and easier to learn communication system. A holistic system able to convey a combinatorial meaning space with dimension n requires m^n words, whereas a compositional system requires only $n \times m$ words to be distributed (m is the number of meanings per dimension of the meaning space). Following Eq. (6), time of convergence for a compositional system is faster than for a holistic system. This explanation has been supported by mathematical analysis (Vogt 2007a).

However, in a recent yet unpublished study by Vogt, in which the model was reimplemented to improve computational efficiency, the transition between phases I and II of the time of convergence did not occur. Instead, as in Baronchelli et al. (2006) and Kaplan (2005), only one power law governing time of convergence was observed, because compositionality had already fully emerged in smaller populations. Although the reasons are not yet fully understood, preliminary analyses of both models indicate that there was a crucial implementation difference between them. In the old model agents were able to generalize over synonyms of holistic elements. For example, when an agent had learned a holistic word for *red square* and later also associated the same word with *blue circle*, this holistic word could, because of the implementation, also erroneously signify *red circle* and *blue square*. This overgeneralization led to an increase in the amount of ambiguity in the language, which in turn reduced the accuracy of the communication system. The reimplementation did not allow for such ambiguities to occur, leading to an improved communicative accuracy and the emergence of compositionality already in small populations.

These differences suggest that, in models in which it is hard to achieve a high level of communicative success, small populations prefer holistic systems to compositional ones, whereas for larger populations compositionality is preferred. In idealized models in which ambiguities are easily dampened (such as in the new implementation), communicative success rates of 100% are obtained (e.g., Kaplan 2005), and compositionality emerges and is preferred over holistic models, even in small populations. However, systems that are cognitively more realistic (e.g., because they are implemented on robots or use a more plausible statistical learning mechanism) tend not to achieve a 100% communicative success rate, typically because of ambiguities arising in the agent's linguistic competence (e.g., Smith 2003; Steels et al. 2002; Vogt 2006; Vogt and Coumans 2003). In the model of Vogt (2007a) ambiguities emerged through an unrealistic generalization mechanism, but it could well be that a more realistic model (e.g., one based on a statistical learning mechanism) might yield qualitatively similar results. Further research is under way to investigate whether this turns out to be the case. Because empirical data suggest that languages spoken in larger communities tend to be more compositional (Levinson 2006; Wray and Grace 2007), this is not unlikely.

Another interesting example of an ABCM that investigates demographic effects on language formation is the study by Steels and McIntyre (1999), in which 20 agents were spatially distributed in three clusters on a two-dimensional grid. The probability that two agents engaged in a language game was inversely proportional to the distance between the two agents. The language game model (aka the naming game model) was similar to the one described in the previous section. So, initially, the agents had no communication system, but they developed their individual lexicons by playing language games.

As in the earlier example, 100% convergence was readily achieved when the agents were equally likely to communicate with each other (i.e., when they were not spatially distributed). When the agents were spatially distributed in three clusters, 100% convergence was obtained only within clusters, although a near 100% convergence was achieved for intercluster communication. Analysis of the emerging lexicons revealed that each cluster had developed its own communication system, which its agents preferred to use. Although some words were used preferentially by all clusters, most words were specific to one cluster. Yet the agents did acquire the words preferred in the other clusters, so that they could understand the agents from other clusters, which explains the high convergence for intercluster communication.

In this first simulation Steels and McIntyre (1999) allowed intercluster communication from the start, but in a second simulation from the same study they allowed only intercluster communication after each cluster had settled on a common communication system, thus simulating language contact. In this second study, Steels and McIntyre showed that, after intercluster communication was allowed, initially a rapid increase in bilingualism emerged, followed by a gradual mixing of the languages. When intercluster communication continued, Steels and McIntyre observed an evolution toward complete coherence. During all this time (except for the period just after the intercluster communication was allowed), the communicative success of the whole system (intracluster and intercluster communication) remained high.

The simulations discussed in this section did not incorporate any empirical data (or empirically derived model) but instead were used to investigate what happens to the language that evolves under various demographic conditions. This has allowed the researchers to discover the conditions under which certain linguistic phenomena can or cannot evolve. The advantage of this approach is that the space of possible structures can be searched structurally without the need to incorporate empirical demographic data that may be hard to find in the literature, may be difficult to translate into the model, or may make the simulation computationally too complex. The vast majority of ABCMs used to investigate the interaction between language evolution and demography have not incorporated empirical data, even though this is desirable. It is not uncommon, however, to compare results of ABCMs qualitatively with empirical data. For instance, Vogt's (2007a) finding that compositionality can emerge more substantially in larger communities than in smaller ones is consistent with findings from languages across the world (Levinson 2006; Wray and Grace 2007). However, the model would be stronger if it could, for instance, predict rather precise quantitative distributions of languages that have various levels of compositionality based on population size such that it is consistent with empirical findings.

Demographic Effects with Empirical Data. Studies that do not incorporate empirical data of demography can investigate only the boundary conditions under which certain linguistic phenomena occur in the computer model. However, to verify a model reliably, the model has to be able to explain the occurrence (evolutionary or otherwise) of real linguistic phenomena based on real demographic data. This, in turn, would improve our insights into the mechanisms underlying such phenomena.

An example of an ABCM that has used empirical demographic data is the model of Briscoe (2000, 2002). He used an ABCM in which the agents were defined with a language acquisition device (LAD) based on a set of possible grammars (i.e., a universal grammar, UG). The LAD was defined as a Bayesian learner who learned by calculating the probabilities that a grammar would be learned given (1) the data triggered by the interaction with other agents (who would use a particular grammar) and (2) a prior bias to learn that grammar as specified by the innate UG. In addition, agents had an age that specified whether they would be learners or nonlearners (adults) and an individual measure of their communicative success rate. Adult agents could reproduce (selection of parents was based on their communicative success rate), and the distribution of prior biases (i.e., UG) in their offspring was initialized by recombining and mutating the genetic material inherited from the parents.

The agents were defined such that the population would grow from an initial state following an S-shaped logistic function. One of Briscoe's experiments aimed to simulate the creolization process of Hawaiian creoles in which different language groups came together and initially developed a pidgin language, followed by the emergence of a more grammatical creole language, as described, for example, by Bickerton (1984). It is beyond the scope of this paper to explain Briscoe's rather complicated model in detail. Suffice it to say that agents were initialized with their prior biases based on genetic evolution and then were triggered to learn one of the possible grammars during the interactions with other agents and using the Bayesian learning mechanism.

The initial distribution of languages in the population was loosely based on demographic information concerning the initial situation of language contact in Hawaii. Briscoe was interested in the period it took for a creole language to emerge and found that his model could explain the catastrophic emergence of creoles from initially pidgin languages in just two generations, as proposed by Roberts (1998). Moreover, Briscoe found that the demographic composition of the initial population of pidgin speakers was important and that the (external) language was adapting to improve learnability within the population (cf. Kirby et al. 2007; Vogt 2007b). Although the use of empirical demographic information was "rather sketchy" and the results were "by no means conclusive" (Briscoe 2000), this example illustrates how an ABCM can be used successfully to simulate empirically studied phenomena such as creolization in Hawaii. The interactions between the genetic evolution of the prior biases, the Bayesian learning mechanisms, the aging of agents, and the reproduction dynamics are so highly complex that it would be extremely hard to capture them in an analytical model or even an ABAM. This type of modeling not only informs us about how demographic processes affect language change but also allows us to investigate the cognitive factors that underlie human language evolution.

Another example of an agent-based-like model is that of Parisi et al. (2008). They used cellular automata to simulate the spread of farming and languages in the European region of the Indo-European language area. Although technically the simulation that Parisi and co-workers used is not an ABCM, I classify it as such because it has a similar level of complexity (see Table 3). What Parisi's group did was to divide Europe into grid cells of 70 square kilometers and define characteristics of each individual cell based on empirical data as much as possible (mostly geographic data). Starting with a small population in one cell in southwest Anatolia and by specifying a population growth within a cell, the carrying capacity of the cell beyond which part of the population would migrate, the likelihood of individuals migrating to a particular neighboring cell, and a few other features, Parisi et al. (2008) were able to simulate the spread of farming such that their results appeared to be in agreement with archaeological data from Neolithic sediments and related models (Ammerman and Cavalli-Sforza 1984; Renfrew 1987; Semino et al. 1996).

From the time it took for certain areas to become populated and the assumption that languages have a certain amount of change over time, Parisi et al. (2008) reconstructed a language tree that to a limited extent resembles the European branch of the Indo-European language tree. Thus their model shows how language families could have spread over Europe based on a demic account in a way that is similar to what is believed to be the way that farming spread. They continued, less convincingly, to expand their model to accommodate a cultural aspect of language transmission (language changes could propagate backward). Although the model can be critiqued on several points (it is beyond the scope of this paper to do so), the method used is a promising one and could be used as a starting point for further studies.

Discussion

In this paper I have reviewed the formal models used to investigate the interaction between demographic processes and language evolution in its broadest sense (i.e., including language origins, early language evolution, language change, and language contact on various linguistic phenomena). The formal models have been classified in three ways: analytical models, agent-based analytical models, and agent-based cognitive models. I have shown that these three classes of models differ in the granularity of detail with which the systems are modeled in terms of demography, linguistics, and sociocognitive mechanisms. Analytical models allow for only a coarse description by having a small number of variables and parameters. ABAMs typically increase the number of variables and parameters by a factor equal to the number of agents in the model. Finally, ABCMs can be even more complex by a potentially large factor of the order of the complexity of the language times the complexity of the cognitive processes underlying language processing.

The choice for selecting which method one should use in studying demographic effects on language evolution depends on the purpose of one's study. Analytical models are particularly suited to investigations of the dynamics of the evolution of and competition between languages based on high-level hypotheses regarding demographic processes and social dynamics. ABAMs are useful for studying the complex interactions between agents defined mathematically in terms of high-level cognitive processes, such as the probability of learning a specific language or grammar. ABCMs are favored when one wishes to investigate the dynamics sprouting from the complex interactions between lower level cognitive processes and social interactions among agents.

The choice could also be weighted by the various advantages and disadvantages of the approaches. The advantage of analytical models is that, because of their low complexity, it is often relatively straightforward to incorporate demographic models based on empirical data. Moreover, analytical models are relatively easy to understand because of their transparency. The disadvantage is that, also because of their low complexity, such models have little explanatory power regarding specific language structures that have evolved.

The advantage of agent-based models is that they provide a significant step forward to more realism, but at the cost of increasing obscurity resulting from increased complexity of the models and at the cost of increasing computational complexity. Nevertheless, such models allow one to investigate social and, in the case of ABCMs, cognitive factors that influence language evolution. Moreover, ABCMs are most suited to investigating the effect that demographic dynamics have on the evolution of languages (i.e., on the emergence and change of linguistic systems).

To assess the realism of sociocognitive models on language evolution, we need validations using demographically realistic simulations. One could argue that a demographically realistic sociocognitive model would be more realistic if its resulting evolution correlated closely with empirically observed linguistic phenomena (Vogt and de Boer 2010). The Baxter et al. (2009) study is a good example. Even though their model does not capture the intricacies of the various sociocognitive mechanisms involved in language processing, Baxter and co-workers suggest that Trudgill's (2004) theory on language change, on which Baxter's model is based, is at least in part correct.

Agent-based modeling of language evolution is a relatively new method and has until recently hardly been used to validate theories based on empirical data [but see, e.g., de Boer (2001) and Nettle (1999b) for early counterexamples]. It took the scientific community quite long to develop models that work sufficiently well for such endeavors to take place. Although an increasing number of studies incorporate empirical data of one kind or another into their models, there is still a long way to go before computer models can convincingly demonstrate the mechanisms underlying language evolution (Vogt and de Boer 2010). The empirical data that need to be incorporated into the models include data from biology, linguistics, psychology, archaeology, and other fields. The studies by Briscoe (2000) and Parisi et al. (2008) are good examples of investigations that have used empirical demographic data to study the evolution of linguistic phenomena in ABCMs.

ABCMs offer an excellent tool for combining all these types of data. However, this is hard to achieve. Typically, the models have been used to investigate conditions under which certain phenomena are observed by searching a large space of possibilities in a certain dimension (e.g., population size). When empirical data are incorporated, they are typically from only one domain, such as linguistics or demography. There are various reasons why empirical data have been incorporated so infrequently. Three of these reasons, which have already been touched on, are the following:

1. Empirical data may be hard to find or are absent. For instance, a precise reconstruction of how *Homo sapiens* populated the world is not available.

2. The empirical data may be hard to translate such that they can be incorporated into a model. The models, even ABCMs, are necessarily crude simplifications and abstractions from reality. Although empirical data can be explained or characterized in abstractions, many theories of language evolution are descriptive rather than formal. These theories often form the basis from which the formal models are abstracted. In other words, the models are abstractions of theories, which are themselves (abstract) descriptions based on empirical data. Hence the distance between model and data is often larger than that between theory and data. Moreover, empirical data are obtained by observing the result of interactions between the real sociocognitive mechanisms, whereas the abstract models capture only an extremely small aspect of these mechanisms. A proper comparison is therefore often infeasible. Some of these problems can be overcome by designing empirical studies that capture those data that we can compare with the models [see also Vogt and de Boer (2010)].

3. Incorporation of empirical data can make the computer simulation intractable. Even though computers are becoming increasingly fast, running an agent-based model with a few thousand agents that develop highly simplified linguistic phenomena may take months of processing. For instance, the study by Vogt (2007a) took over a month of processing on a cluster of 10 standard PCs and included only 140 simulations of 5 generations with populations varying in size up to 300 agents who could learn only two-word sentences describing colored shapes. (The slow performance was partly due to the ineffective computer program, but the more effective program developed for the followup study used almost a month to process 200 simulations of 10 generations with groups of up to 1,000 agents on the same machines.) It is to be expected that with the ongoing increase in computational power, this problem will resolve itself to some extent, so computers become ever more suitable for dealing with the complex processes of language evolution in agent-based models.

Because studying the complex sociocognitive and demographic processes using ABCMs is becoming more feasible, the question emerges, What are the logical next steps? One possibility is to simulate the emergence of Nicaraguan Sign Language (Kegl and Iwata 1989), which is a well-documented case in both demography and the evolution of a full-grown sign language. Such data form an ideal basis to study the cultural evolution of language and have similarities to the creolization processes of language, so Briscoe's studies could be a starting point, as are those by Baxter et al. (2009), Minett and Wang (2008), and Vogt (2005b). There is also a possibility of studying the relation between group sizes and language evolution in order to further investigate the findings of Vogt (2007a). For instance, one could try to incorporate archaeological data about population growth to study how this would effect language change. In addition, one could incorporate empirical quantitative data on social network structures to further improve the realism of the computer simulations.

Another avenue would be to incorporate demographic data on age structures in populations in the computer simulations. It has been suggested that young children may have had a large influence on the formation of grammatical structures in language (e.g., Bickerton 1984; Senghas et al. 2004; Vogt 2005b). Moreover, language use tends to change over different age groups (Boberg 2004). So, it is interesting to investigate how different age distributions of a population affect language change over a longer period of time. A quick survey on the Internet did not reveal that such studies exist, although Briscoe's studies do incorporate an age structure, but he did not back this with empirical data.

Finally, the grand challenge would be to further develop ABCMs along the lines of Parisi et al. (2008) in order to study how languages have evolved and occupied, together with humans, the world. With an increasing amount of archaeological data concerning the demographic patterns of the spread of *Homo sapiens* in the world (Klein 2000; Mellars 2005), theories could be tested about whether such dynamics would, indeed, lead to language trees that are similar to those found in reality. Although it is highly unlikely that investigators would obtain an exact copy of the world's languages, the statistical distribution of languages, their diversity, and rate of change should closely match those found in the world.

Conclusions

Three formal methods for studying the effects that demographic processes can have on language evolution have been reviewed and characterized. These three methods (analytical modeling, agent-based analytical modeling, and agent-based cognitive modeling) differ primarily in the complexity of the processes that the model can capture; the analytical models are the least complex, and the agent-based cognitive models are the most complex. The purpose of this paper was to provide readers interested in studying the interactions between demography and language evolution with a guideline for selecting one of these modeling approaches on the basis of their characteristics, advantages, and disadvantages.

Whatever the language evolution phenomenon is that one wants to test, models that take empirical demographic data as a way to set up computer simulations would be ideal benchmark models to validate theories of language evolution, provided that their outcome could be compared with empirical data on that linguistic phenomenon. These benchmark studies could be centered around the grand challenge of simulating the spread of languages across the world, but given the complexity of such models, setting up benchmarks on smaller challenges (e.g., cf. Baxter et al. 2009; Briscoe 2000) would, at least for the near future, be more appropriate and easier to achieve.

Although analytical models can help us to gain insights into the high-level conditions under which certain linguistic phenomena may occur, they fail to capture the vast complexity of human language processing that affects language evolution in complex societies. Agent-based analytical models are an important step forward, because they can capture the complexity that analytical models put in populations in a single agent so that one can study the effects of social interactions on language evolution. However, the human cognitive apparatus is far more complex than can be captured in a mathematical formula. Therefore agent-based cognitive modeling, although still limited, is an essential method for researching the sociocognitive interactions supposed to underlie the evolution of human language and languages.

Acknowledgments I thank the participants of the workshop on "Demographic Processes and Cultural Change," organized by the Arts and Humanities Research Council (AHRC) Center for the Evolution of Cultural Diversity, that was held at Missenden Abbey in September 2008, in particular the organizers, Stephen Shennan and James Steele, for creating an inspiring environment that led to the writing of this paper. I also thank the Netherlands Organization for Scientific Research (NWO) for supporting me through a Vidi grant.

Received 13 January 2009; revision accepted for publication 16 April 2009.

Literature Cited

Abrams, D. M., and S. H. Strogatz. 2003. Modeling the dynamics of language death. *Nature* 424:900.

Ammerman, A., and L. Cavalli-Sforza. 1984. *Neolithic Transition and the Genetics of Population in Europe.* Princeton, NJ: Princeton University Press.

Baronchelli, A., M. Felici, E. Caglioti et al. 2006. Sharp transition towards shared lexicon in multi-agent systems. *J. Stat. Mechanics* P06014.

Batali, J. 1998. Computational simulations of the emergence of grammar. In *Approaches to the Evolution of Language*, J. R. Hurford, M. Studdert-Kennedy, and C. Knight, eds. Cambridge, U.K.: Cambridge University Press, 405–426.

Baxter, G. J., R. A. Blythe, W. Croft et al. 2006. Utterance selection model of language change. *Phys. Rev. E* 73:046118.

Baxter, G. J., R. A. Blythe, W. Croft et al. 2009. Modeling language change: An evaluation of Trudgill's theory of the emergence of New Zealand English. *Lang. Var. Change* 21:257–296.

Bickerton, D. 1984. The language bioprogram hypothesis. *Behav. Brain Sci.* 7:173–212.

Boberg, C. 2004. Real and apparent time in language change: Late adoption of changes in Montreal English. *Am. Speech* 79:250–269.

Boyd, R., and P. J. Richerson. 1985. *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.

Briscoe, E. J. 2000. Evolutionary perspectives on diachronic syntax. In *Diachronic Syntax: Models and Mechanisms*, S. Pintzuk, G. Tsoulas, and A. Warner, eds. Oxford, U.K.: Oxford University Press, 75–108.

Briscoe, E. J. 2002. Grammatical acquisition and linguistic selection. In *Linguistic Evolution Through Language Acquisition: Formal and Computational Models*, E. J. Briscoe, ed. Cambridge, U.K.: Cambridge University Press, 255–300.

Cangelosi, A., and D. Parisi, eds. 2002. *Simulating the Evolution of Language*. London: Springer.

Cavalli-Sforza, L. L., and M. W. Feldman. 1981. *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton, NJ: Princeton University Press.

Christiansen, M. H., and S. Kirby, eds. 2003. *Language Evolution*. Oxford, U.K.: Oxford University Press.

Dall'Asta, L., A. Baronchelli, A. Barrat et al. 2006. Nonequilibrium dynamics of language games on complex networks. *Phys. Rev. E* 74:036105.

de Boer, B. 2001. *The Origins of Vowel Systems*. Oxford, U.K.: Oxford University Press.

de Boer, B. 2005. Evolution of speech and its acquisition. *Adapt. Behav.* 13(4):281–292.

Gong, T., J. W. Minett, and W. S. Y. Wang. 2008. Exploring social structure effect on language evolution based on a computational model. *Connection Sci.* 20(2–3):135–153.

Griffiths, T. L., M. L. Kalish, and S. Lewandowsky. 2008. Theoretical and empirical evidence for the impact of inductive biases on cultural evolution. *Philos. Trans. R. Soc. Lond. B* 363(1509):3503–3514.

Kandler, A. 2009. Demography and language competition. *Hum. Biol.* 81(2–3):181–210 (this issue).

Kandler, A., and J. Steele. 2008. Ecological models of language competition. *Biol. Theor.* 3:164–173.

Kaplan, F. 2000. *L'émergence d'un lexique dans une population d'agent autonomes*. Ph.D. dissertation, Laboratoire d'informatique de Paris 6.

Kaplan, F. 2005. Simple models of distributed coordination. *Connect. Sci.* 17(3–4):249–270.

Ke, J.-Y., T. Gong, and W. S.-Y. Wang. 2008. Language change and social networks. *Commun. Computat. Phys.* 3(4):935–949.

Kegl, J., and G. Iwata. 1989. Lenguaje de Signos Nicaraguense: A pidgin sheds light on the "creole?" In *Proceedings of the Fourth Annual Meeting of the Pacific Linguistics Conference*, R. Carlson, S. DeLancey, S. Gilden et al., eds. Eugene: University of Oregon, Department of Linguistics, 266–294.

Kirby, S. 2001. Spontaneous evolution of linguistic structure: An iterated learning model of the emergence of regularity and irregularity. *IEEE Trans. Evol. Computat.* 5(2):102–110.

Kirby, S. 2002. Natural language from artificial life. *Artif. Life* 8(3):185–215.

Kirby, S., M. Dowman, and T. Griffiths. 2007. Innateness and culture in the evolution of language. *Proc. Natl. Acad. Sci. USA* 104(12):5241–5245.

Kirby, S., K. Smith, and H. Brighton. 2004. From UG to universals: Linguistic adaptation through iterated learning. *Stud. Lang.* 28(3):587–607.

Klein, R. G. 2000. Archeology and the evolution of human behavior. *Evol. Anthropol.* 9:17–36.

Komarova, N. L., P. Niyogi, and M. A. Nowak. 2001. The evolutionary dynamics of grammar acquisition. *J. Theor. Biol.* 209(1):43–59.

Levinson, S. C. 2006. Introduction: Evolution of culture in a microcosm. In *Evolution and Culture: A Fyssen Foundation Symposium*, S. Levinson and P. Jaisson, eds. Cambridge, MA: MIT Press, 1–41.

Lightfoot, D. 1999. *The Development of Language: Acquisition, Change, and Evolution*. Oxford, U.K.: Blackwell.

Livingstone, D., and C. Fyfe. 2000. Modeling language-physiology coevolution. In *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*, J. R. Hurford, C. Knight, and M. Studdert-Kennedy, eds. Cambridge, U.K.: Cambridge University Press, 199–215.

Mellars, P. 2005. The impossible coincidence: A single-species model for the origins of modern human behavior in Europe. *Evol. Anthropol.* 14(1):12–27.

Minett, J. W., and W. S. Wang. 2008. Modeling endangered languages: The effects of bilingualism and social structure. *Lingua* 118(1):19–45.

Nettle, D. 1999a. Is the rate of linguistic change constant? *Lingua* 108:119–136.

Nettle, D. 1999b. *Linguistic Diversity*. Oxford, U.K.: Oxford University Press.

Nettle, D. 1999c. Using social impact theory to simulate language change. *Lingua* 108:95–117.

Niyogi, P., and R. C. Berwick. 1996. A language learning model for finite parameter spaces. *Cognition* 61(1–2):161–193.

Nowak, M. A., N. L. Komarova, and P. Niyogi. 2001. Evolution of universal grammar. *Science* 291:114–118.

Oliphant, M. 1996. The dilemma of saussurean communication. *Biosystems* 1–2(37):31–38.

Oudeyer, P.-Y. 2005. How phonological structures can be culturally selected for learnability. *Adapt. Behav.* 13:269–280.

Parisi, D., F. Antinucci, F. Natale et al. 2008. Simulating the expansion of farming and the differentiation of European languages. In *Origin and Evolution of Languages: Approaches, Models, Paradigms*, B. Laks, ed. London: Equinox, 234–258.

Patriarca, M., and T. Leppänen. 2004. Modeling language competition. *Physica A* 338(1–2):296–299.

Renfrew, C. 1987. *Archaeology and Language: The Puzzle of Indo-European Origins*. London: Jonathan Cape.

Roberts, S. 1998. The role of diffusion in the genesis of Hawaiian creole. *Language* 74(1):1–39.

Semino, O., G. Passarino, A. Brega et al. 1996. A view of the Neolithic demic diffusion in Europe through two Y-chromosome-specific markers. *Am. J. Hum. Genet.* 59:964–968.

Senghas, A., S. Kita, and A. Özyürek. 2004. Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science* 305(5691):1779–1782.

Smith, A. D. M. 2003. Intelligent meaning creation in a clumpy world helps communication. *Artif. Life* 9(2):559–574.

Smith, K., H. Brighton, and S. Kirby. 2003. Complex systems in language evolution: The cultural emergence of compositional structure. *Adv. Complex Syst.* 6(4):537–558.

Steels, L. 1996. Emergent adaptive lexicons. In *From Animals to Animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior*, P. Maes, M. Mataric, J.-A. Meyer et al., eds. Cambridge, MA: MIT Press, 562–567.

Steels, L. 2005. The emergence and evolution of linguistic structure: From lexical to grammatical communication systems. *Connect. Sci.* 17(3–4):213–230.

Steels, L., F. Kaplan, A. McIntyre et al. 2002. Crucial factors in the origins of word-meaning. In *The Transition to Language*, A. Wray, ed. Oxford, U.K.: Oxford University Press, 252–271.

Steels, L., and A. McIntyre. 1999. Spatially distributed naming games. *Adv. Complex Syst.* 1(4):301–323.

Trudgill, P. 2004. *New-Dialect Formation: The Inevitability of Colonial Englishes*. Edinburgh, U.K.: Edinburgh University Press.

Vogt, P. 2005a. The emergence of compositional structures in perceptually grounded language games. *Artif. Intel.* 167(1–2):206–242.

Vogt, P. 2005b. On the acquisition and evolution of compositional languages: Sparse input and the productive creativity of children. *Adapt. Behav.* 13(4):325–346.

Vogt, P. 2006. Language evolution and robotics: Issues in symbol grounding and language acquisition. In *Artificial Cognition Systems*, A. Loula, R. Gudwin, and J. Queiroz, eds. Hershey, PA: Idea Group, 176–209.

Vogt, P. 2007a. Group size effects on the evolution of compositional languages. In *Advances in Artificial Life: Proceedings of ECAL-2007 (LNAI-4684)*, F. Costa e Almeida, L. M. Rocha, E. Costa et al., eds. Berlin: Springer, 405–414.

Vogt, P. 2007b. Variation, competition, and selection in the self-organization of compositionality. In *The Mind, the Body, and the World: Psychology After Cognitivism?* B. Wallace, A. Ross, J. B. Davies et al., eds. Exeter, U.K.: Imprint Academic, 233–256.

Vogt, P., and H. Coumans. 2003. Investigating social interaction strategies for bootstrapping lexicon development. *J. Artif. Soc. Social Simul.* 6(1). http://jasss.soc.surrey.ac.uk/6/1/4.html.

Vogt, P., and B. de Boer. 2010. Language evolution: Computer models for empirical data. *Adapt. Behav.* (in press).

Wray, A., and G. W. Grace. 2007. The consequences of talking to strangers: Evolutionary corollaries of sociocultural influences on linguistic form. *Lingua* 117(3):543–578.