*Evolutionary Ecology Research*, 2009, **11**: 433–445

# Simplicity from complexity: emergence of cohesion in the evolutionary dynamics of grammar networks

Yoosook Lee[1], Travis C. Collier[1], Charles E. Taylor[1],
and Reza Olfati-Saber[2]

[1]*Department of Ecology and Evolutionary Biology, UCLA, Los Angeles, California and*
[2]*Thayer School of Engineering, Dartmouth College, Hanover, New Hampshire, USA*

## ABSTRACT

**Question:** Under what conditions will a group of adaptive, communicating agents converge on a single language, and when will there instead be a multiplicity of languages? This is sometimes termed the 'Tower of Babel problem', and is similar to those addressed in Slobodkin's book on simplicity and complexity.

**Features of the model:** Agents learn to speak a language from those around them locally, or from those with whom they communicate. Learning fidelity can vary, as can connectedness. Populations of learning agents are simulated and the equilibrium properties deduced.

**Range of key variables:** Populations of agents are arranged in a regular ring lattice, with varying degrees of connectivity. Languages are not all equally related, as studied in the past, but rather some languages are more closely related than others.

**Conclusions:** If a grammar has limited mutational transitions to other grammars, one expects three, rather than two, distinctive grammar distributions at equilibrium. In line with previous studies, our model exhibits a *bifurcation* from the *symmetric state*, where all grammars are in equal frequency, to an *asymmetric state* under dense grammar networks. However, different patterns of convergence, including a *weak cohesion* phase where the distribution of grammar frequencies is roughly Gaussian and centred on the most frequent grammar, emerge in more sparse grammar networks. Weak cohesion is of particular interest from an adaptation standpoint, since there is both linguistic coherence and standing heritable variation.

*Keywords*: evolutionary dynamics, grammar networks, language evolution.

## INTRODUCTION

In *Simplicity and Complexity in Games of the Intellect*, Larry Slobodkin (1992) recounted how a lifetime of research on seemingly simple organisms like polyps led him to consider just what is meant by 'simple' in the context of biology and human affairs. After discussing the

---

Correspondence: Y. Lee, Department of Ecology and Evolutionary Biology, UCLA, 621 Charles Young Dr. South (LS 3109), Los Angeles, CA 90095, USA. e-mail: yoslee@ucdavis.edu

concept in art, religion, and science, he concluded: 'I had hoped that I could develop a single more or less formal theory of simplicity or, better yet, find one ready made in, perhaps, the philosophical literature. I found no general theory' (Slobodkin, 1992, p. 223).

In fact, such a theory had been introduced by Jorma Rissanen in 1978 (see Rissanen and Ristad, 1994), termed 'minimum description length' (MDL). It has become a cornerstone of current thinking in computational linguistics and learning theory. Roughly speaking, MDL defines a simpler theory as one that requires less information to describe its domain, using one from a set of allowable languages. It is not clear that this theory is applicable to cnidarians or human affairs – Larry's principal domains of interest, if he can be said to have such. It is, however, clearly applicable to many of the systems being studied in artificial life, in particular to adaptive networks of sensor arrays such as those we have been studying. The principal goal of our research is to develop networks of agents that can sense their environment, communicate with one another about it, arrive at a general understanding, and then communicate that understanding to a human observer when appropriate. This is an ambitious goal, and one in which language plays a central role. We wish the languages of the array to be expressive (hence capable of complex semantics) but at the same time to be both groundable and learnable (requiring simplicity). We have made some progress towards our goal (see Stabler *et al.*, 2003; Kobele *et al.*, 2004; Ali *et al.*, 2007; Escobar *et al.*, 2007).

A recurring problem with this work has been coherence in the language(s) that is (are) being learned by the network. Imagine, for example, that the network hears a type of event (e.g. new bird song) in one part of the array and assigns a symbol to that (e.g. species-x). At about the same time another part of the array hears the same song and assigns a different symbol (e.g. species-y) to the song. How are these to be reconciled? (There is a similarity to the classic problem of modal logic, of recognizing that the morning star is the same as the evening star.) We would like for all agents in the array to speak a single, coherent, mutually intelligible language instead of each agent going off on its own, contributing to a mini 'Tower of Babel'. For the last decade, great progress has been made in the field of robotics to investigate the social and/or cognitive factors contributing to the emergence of coherent communication among autonomous agents (Arita and Taylor, 1996; Steels 2001, 2006; Marocco *et al.*, 2003; Collier and Taylor, 2004; Lee *et al.*, 2004; Marocco and Nolfi, 2006). The dynamics of language evolution provides insight into convergence to a common understanding where distributed learning is a goal.

In this study, we address the dynamics of a population of language learners as a less abstract instantiation of the general case of agents learning from one another. Grammars may be viewed as any state or knowledge base learned or derived from other agents where commonality is advantageous.

Recent studies applying a biological perspective to the evolution of linguistic convergence and diversity have shown promising results (Hashimoto and Ikegami, 1995, 1996; Steels, 1996; Kirby, 2001; Komarova *et al.*, 2001; Niyogi, 2006). In particular, Komarova, Nowak, and Niyogi have adopted a model based on Eigen and Schuster's molecular evolution work (Eigen and Schuster, 1979; Fontana and Schuster, 1987; Eigen *et al.*, 1989) for describing the dynamics of language evolution (Komarova and Nowak, 2001; Komarova *et al.*, 2001; Niyogi, 2006).

Parameters such as the maximum error rate a communication channel can withstand before the agents lose their ability to communicate in a coherent language have been determined for some regimes (Komarova *et al.*, 2001; Nowak *et al.*, 2001). However, most strong results have been based on the convergence dynamics of a population learning languages from a set where each language is equally related to every other language. In contrast, natural languages are organized hierarchically, with language families, subgroups, and even dialects. Naturally, some languages are more closely related than others. The similarity of this language hierarchy with species trees is one of the inspirations for an evolutionary approach in the first place.

In this paper, we define a grammar network to be a graph describing the similarity between grammars. We assume that the grammar network determines the distance of grammars from each other in mutational space as well as their contribution to mutual intelligibility. Such a network, therefore, represents a space where only transitions between grammars that are connected by edges are allowed in a single reproduction/learning step. Preliminary results have shown that different grammar networks alone can qualitatively change the dynamics of convergence (Olfati-Saber and Murray, 2004; Lee *et al.*, 2005; Olfati-Saber, 2005). Since our setting does not have a limit on population size, our result may not be directly comparable with the independent study by Matsen and Nowak (2004), who recently explored language convergence conditions on a 'nearly-regular' language network.

Here we explore the population dynamics of grammar acquisition using a grammar network structured as a ring lattice. Although we have also explored other network structure such as small-world and random network, we are compelled to focus on ring lattice because it illustrates how grammar network alone can change the dynamics of language convergence.

Two equilibrium states have previously been characterized in terms of language convergence: one is the symmetric state where all grammars exist in equal frequencies, and the other is an asymmetric state where the symmetry breaks and one grammar predominates in the population (Komarova *et al.*, 2001; Nowak *et al.*, 2001).

Our results identify another state we label 'weak cohesion', where a group of structurally similar grammars predominates in the population. The high degree of standing variation characteristic of the weak cohesion state is more suggestive of real-world observations than the other two equilibria previously described. The weak cohesion regime appears desirable for many adaptive systems, since it achieves a high level of cohesion without eliminating too much variation and losing further adaptability.

## Evolutionary dynamics of a population of language learners

Consider a network of grammars with nodes $U = \{G_1, G_2, \ldots, G_n\}$. Let $a_{ij}$ denote the probability that a learner acquires a grammar $G_j$ from a teacher who speaks grammar $G_i$. The matrix $A = [a_{ij}]$ is determined by the set of edges $E$ of a particular graph $\Gamma = (U, E)$. In our setting, $A$ is defined as: $a_{ij} = 1$ if $i = j$, $a_{ij} = a$ if $(i, j)$ is in $E$, $a_{ij} = 0$ if $(i, j)$ is not in $E$. Matrix $A$ represents the interconnection topology of the grammar network.

Each grammar represents a language hypothesis, consisting of a set of rules that an individual deduced from its input. It is possible that two grammars can have a completely different set of rules and yet generate sentences that are somewhat mutually understandable. To prevent potential confusion, let us distinguish between two types of similarities, structural similarity and expressive similarity. Structural similarity is how many grammatical rules or lexical items two grammars share. Expressive similarity relates to the probability that a sentence generated from one grammar is similar to a sentence generated from another grammar. Structural similarity is analogous to genotype similarity, and expressive similarity is analogous to phenotype similarity.

In our setting, all grammars in a network are positioned on a polygon where their positions are indicative of structural similarity. If two grammars are positioned side by side, they share many common rules for generating sentences.

The relationship between grammars is represented as edges in a grammar network specifying which learner grammars each teacher grammar may produce. In a molecular evolution framework, this is the graph defining possible mutational transitions. For our model, the grammar network is completely specified by the matrix of mutual similarities $A$. Note that a mutual similarity $a_{ij}$ is equivalent to the expressive similarity of two language hypotheses. As long as two grammars are connected ($(i, j)$ is in $E$), they have some degree of mutual intelligibility ($a_{ij} > 0$).

Within this setting, each mutation or incorrect language learning step does not necessarily yield a grammar that is structurally similar to its teacher's grammar. Sometimes the learner can deduce a set of rules that are completely different from that of its parent and yet generate sentences that are very close to its input. Thus, the grammar network defines the space that an individual explores while learning language.

Let $x_i$ denote the proportion of an infinitely large population speaking grammar $G_i$ with $n$ possible grammars existing. We assume that each individual uses only one grammar, thus we have $\Sigma\, x_j = 1$.

The fitness of individuals with grammar $G_i$ is $f_i = f_0 + \Sigma_j a_{ij} x_j$, where $f_0$ is the base fitness, which does not depend on the language, and $\Sigma_j a_{ij} x_j$ is the language contribution to fitness. Note that this fitness equation is frequency dependent. The fitness increases as the product of the proportion of people who use grammar $G_j$ and its similarity with $G_i(a_{ij})$. As more people use the same or more similar grammars, the fitness increases.

The evolutionary dynamics of this population is described as the change in the proportion of the population speaking $G_i$ in the form of

$$\dot{x}_i = \sum_j x_j f_j q_{ji} - \varphi x_i, \qquad 1 \le j \le n, \tag{1}$$

where $\varphi = \Sigma_i x_i f_i$ is the average fitness, and $Q = [q_{ij}]$ is the learning fidelity matrix. The first term ($\Sigma_i x_i f_i q_{ji}$) describes all the individuals learning grammar $G_i$ from teachers speaking all possible grammars including $G_i$ and their fitness contributions. The term $-\varphi x_i$ maintains a constant population size. Equation (1) is known as the replicator-mutator dynamics in the generally accepted framework inspired by Eigen and Schuster (1979). $Q$ is basically a Markov chain transition matrix, which is used as a learning fidelity matrix here and also by Nowak, Komarova (Komarova *et al.*, 2001), and others. This dynamic system can be thought of either as having individuals that produce offspring to replace a randomly chosen individual, or as having individuals that change their grammars by learning a teacher's language. We will use the latter terminology.

The learning model is a relationship between the matrix of mutual similarities $A$ and the transition matrix $Q$, which is defined by

$$q_{ii} = q, \qquad q_{ij} = (1 - q)\frac{a_{ij}}{\sum_{j \ne i} a_{ij}} \text{ for all } i \ne j. \tag{2}$$

The learning fidelity $q$ is the probability that a learner acquires the same grammar as its teacher. A learner has probability $q_{ii}$ of learning a correct grammar $G_i$ from a teacher who speaks $G_i$. If the learner fails to learn the correct grammar, it has probability $q_{ij}$ to learn another grammar $G_j$ and the transition probability is proportional to the relative similarity among other grammars. $Q$ satisfies the condition $\Sigma_j q_{ij} = 1$ for all $i$. The special case of this transition matrix where $a_{ij} = a$ for all $i \neq j$ was analysed by Komarova and colleagues (Komarova and Nowak, 2001; Komarova *et al.*, 2001).

## Grammar networks

Graph theory provides a large number of metrics for describing the topology of grammar networks (Olfati-Saber and Murray, 2004; Olfati-Saber, 2005). The two we will focus on are density and the mean path length. Density $D$ is the measure of the number of links in the graph divided by the number of possible links (Albert and Barabasi, 2002). Mean path length is the mean number of links in the graph that must be traversed to connect any pair of nodes.

For graphs of differing density and mean path length, we expect the dynamics to be different, since a change in $A$ also changes $Q$ as well as fitness. Less than complete graphs make sense from a biological or sociological standpoint, since learning a closely related language is more common than learning a drastically different language.

We systematically analyse the effects of learning fidelity $q$ on the evolutionary dynamics defined in equation (1) for two canonical types of graphs with substantially different structural properties. These are the complete graph and regular ring lattice graphs, depicted in Fig. 1.

### *The complete graph*

The complete graph with all-to-all links is considered by Komarova and colleagues (Komarova *et al.*, 2001; Nowak *et al.*, 2001) because the graph provides a simplified setting where finding an analytic solution is relatively tractable (example shown in Fig. 1A). The adjacency matrix $A$ with similarity parameter $a > 0$ is defined by: $a_{ij} = 1$ if $i = j$ and $a_{ij} = a$ if $i \neq j$.

The total number of links for the complete graph is $n(n - 1)/2$, and the density is 1. In terms of grammar networks, the complete graph is one in which every grammar has equal mutual intelligibility with every other grammar and an individual can change its grammar to any other grammar with a single step evident from mean path length equal to 1.

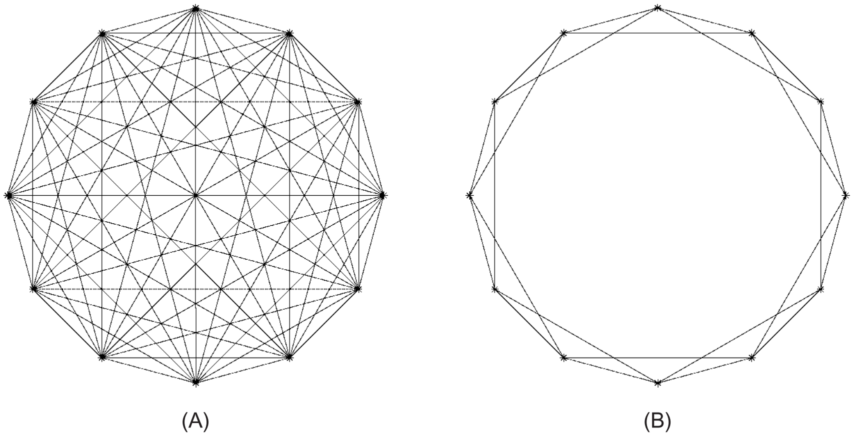

**Fig. 1.** Two types of network structure tested in this paper. Examples are drawn with 12 nodes. (A) The complete graph; (B) a ring lattice with $k = 2$.

### *Regular ring lattice*

A regular ring lattice graph is a cyclic graph with $n$ nodes evenly spaced on a ring and each node linked to its $2k$ nearest neighbours on the ring (example shown in Fig. 1B). This type of graph is also known as a one-dimensional lattice. The adjacency matrix of this graph is denoted by

$$A = I_n + aC(n, k)$$

where $I_n$ denotes an $n \times n$ identity matrix. $C(n, k)$ is a $n \times n$ matrix representing a regular ring lattice, where $c_{ij} = 1$ if node $i$ and $j$ are connected by an edge.

A ring lattice has a total of $nk$ links. The density of this class of graphs is $2nk/n(n-1) = 2k/(n-1)$, which for $2k \ll (n-1)$ is much less than 1. For the special choice of $n = 2k + 1$, one obtains a complete graph with the maximum density of 1.

In a ring lattice network, the mutual similarity can infer some structural similarity in the set of rules. For a learner to acquire a grammar that is very different from that of its teacher, the learner has to undergo many mutational steps. This trend is reflected by the large mean path length, which is proportional to $n = k$ on a ring lattice (Albert and Barabasi, 2002).

## METHODS

Results are obtained from numerical evaluation of the system described by equation (1), referred to as runs, using the fourth-order Runge-Kutta method (Abramowitz and Stegun, 1972). We assume the population is approximately in an equilibrium state when the first and second derivatives for all grammar frequencies are below a predefined threshold of $10^{-12}$.

We obtain the approximate equilibrium frequencies of grammars for a given adjacency matrix $A$ associated with a graph $G$, and a random choice of initial state $x(0)$, which satisfies $\Sigma_i x_i(0) = 1$. From the grammar frequencies we calculate the *effective number of grammars* $n_e$ as a measure of grammar diversity, which is defined as: $n_e = 1/\sum_i x_i^2$ (Crow and Kimura, 1970).

## RESULTS

In Fig. 2 we show the equilibrium frequency of the dominant grammar over a range of $q$ values for a regular lattice graph with fixed $n = 51$ and $a = 0.5$. Each point is the result from a single run, but the $q$ interval ($= 10^{-4}$) is small enough that the points appear as a line in places. The symmetric solution where all grammars exist in equal frequencies of $1/n$ is a stable attractor for $q \leq q_s$. The asymmetric state where one predominant grammar emerges becomes a stable attractor above a critical $q$ value of $q_a$. Figure 2A shows the bi-stability region, where both symmetric and asymmetric solutions are stable, in line with the results analysed by Komarova and colleagues (Komarova *et al.*, 2001; Nowak *et al.*, 2001). In a complete graph, $q_s$ is greater then $q_a$, allowing a bi-stability region where both symmetric and asymmetric solutions can be stable depending on the initial frequency distribution.

The ring lattice graph with $k = 15$ also has symmetric and asymmetric solutions, as shown in Fig. 2B. For the ring lattice network, the symmetric state breaks at $q_s$ well before asymmetric solutions can be stable. The symmetric state is not observed for the ring lattice network with $k = 10$, as shown in Fig. 2C.

The learning fidelity threshold for asymmetric solutions, $q_a$, is highest for the complete graph and it is lower in the networks with smaller $k$ values ($q_a$, $k = 25 = 0.9720 > q_a$, $k = 15 = 0.8280 > q_a$, $k = 10 = 0.7956$).

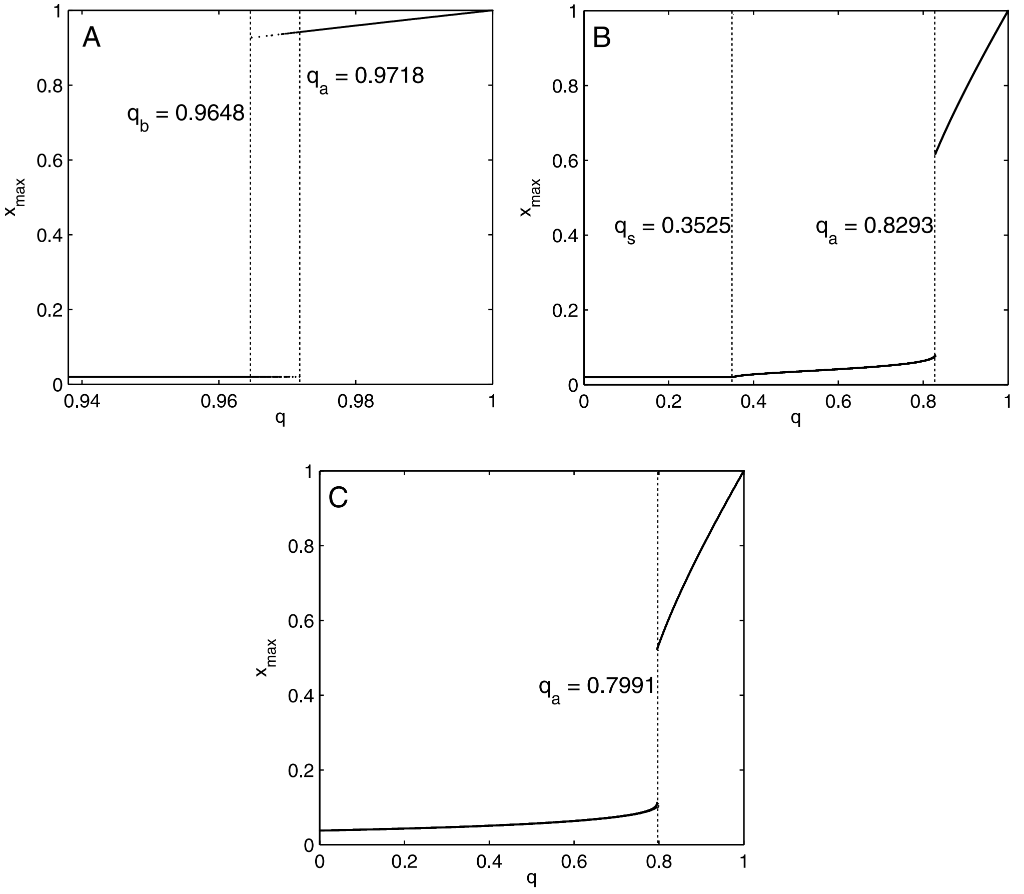

**Fig. 2.** The dominant grammar frequency $x_{\max}$ versus learning fidelity $q$ for a ring lattice with $n = 51$. (A) $k = 25$ (a compete graph), (B) $k = 15$, and (C) $k = 10$. The $q$ interval is $10^{-4}$.

Figure 3 shows the effective number of grammars ($n_e$) for precisely the same runs as Fig. 2 (see Methods section for the definition of $n_e$). In agreement with the results shown in Fig. 2A, Fig. 3A shows two classes of equilibrium states as well as a bi-stability region over a range of learning fidelities $q_a \leq q \leq q_s$ for the complete graph. We calculated the effective number of grammars $n_e$ as a measure of grammar diversity. The $n_e$ is maximum ($= n$) when a population is in a symmetric state, while $n_e$ approaches minimum (~1) when a single dominant grammar emerges. The value of $n_e$ is indicative of the degree of linguistic coherence.

Figure 3B suggests a third class of solutions that occurs at $q$ values between $q_s$ and $q_a$ for ring lattice networks. This class of stable attractors is characterized by a nearly linear decrease in $n_e$ when $q_s \leq q \leq q_a$, as shown in Fig. 3B.

For the regular ring lattice with $k = 10$ (Figs. 2C and 3C), the symmetric state does not exist for $q \geq 0$, as shown in Fig. 3C. The symmetric solution can still be obtained using negative $q$ values, but the interpretation of such values is not obvious. Figure 3C also shows that some level of coherence can be achieved even with a learning fidelity of 0 when $k = 10$.

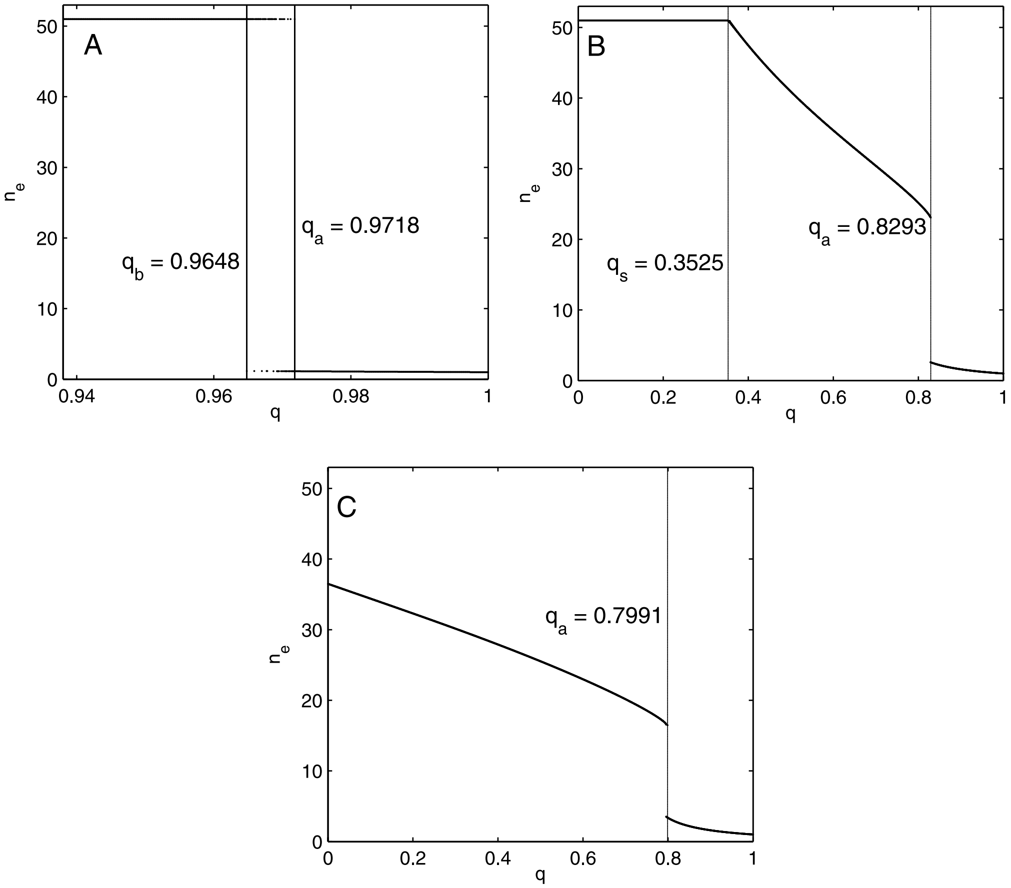

**Fig. 3.** The effective number of grammars $n_e$ versus learning fidelity $q$. (A) $k = 25$ (complete graph), (B) $k = 15$, and (C) $k = 10$.

In the regular ring lattice with $k = 15$ (Figs. 2B and 3B), the frequencies of each grammar $x_i$ at an approximate equilibrium state for different learning fidelities are distributed as shown in Fig. 4. We cut the ring at the point opposite to the dominant grammar and spread it along the x-axis, so that the dominant grammar is always at the centre. If the grammar is positioned close to the dominant, the grammar index is close to the index of the dominant, indicating that they are structurally similar. Figure 4A shows the grammar frequency distribution when $q = 0$ as an example of a symmetric state. If the learning fidelity $q$ is greater than $q_a$, only one grammar dominates, as shown in Fig. 4C. We call this phase *strong cohesion*. When the learning fidelity is between $q_s$ and $q_a$, the grammar frequencies form a smooth curve reminiscent of a Gaussian, as shown in Fig. 4B. We call this phase *weak cohesion*. In this phase, learning fidelity is too low for a single grammar to dominate by faithfully reproducing itself, although structure in grammar space allows for a collection of closely structurally similar grammars to rise in frequency. Since learning errors produce similar grammars to the teacher's grammar, the effective learning fidelity for the group of grammars is higher. This is analogous to the formation of a quasi-species in molecular evolution (Eigen and Schuster, 1979).

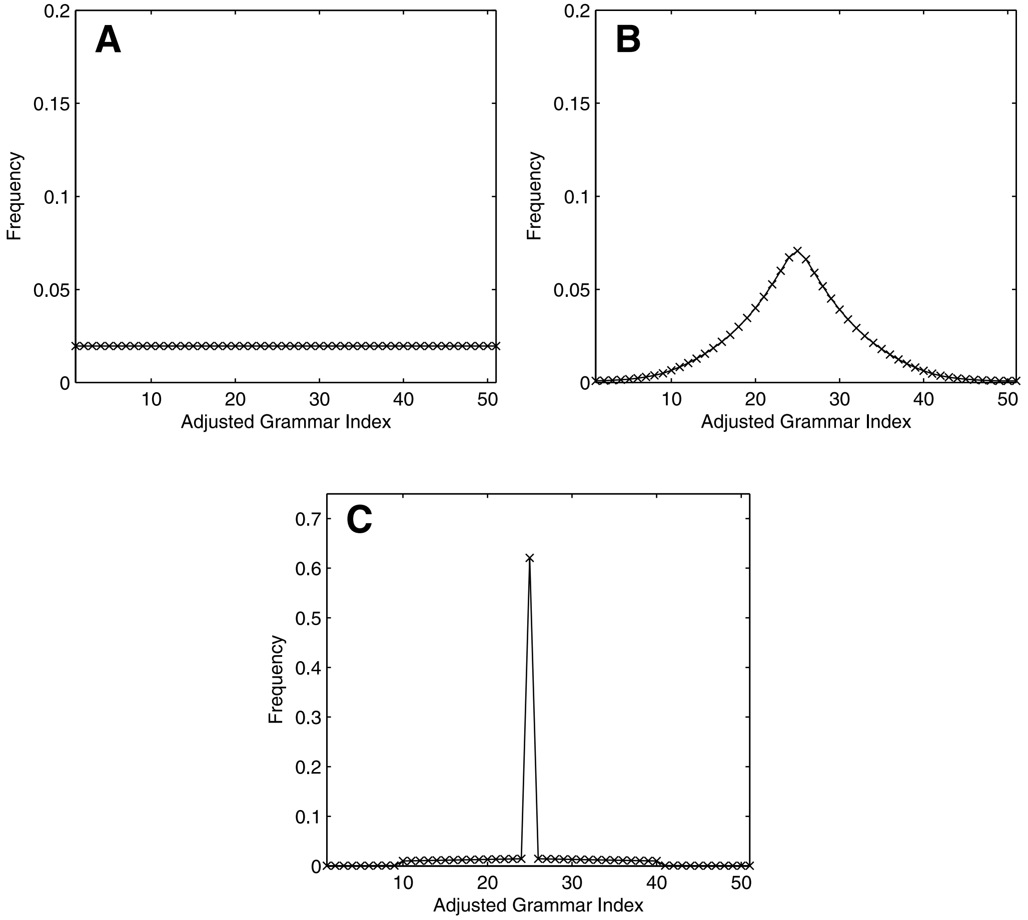

**Fig. 4.** Frequency of all grammars, sorted by index number, of a single run at the steady state for a ring lattice with $n = 51$, $k = 15$, and $a = 0.5$. (A) $q = 0$, (B) $q = 0.82$, and (C) $q = 0.83$.

If a grammar space is structured as a complete graph, high-fidelity learner-driven change, such as the sort exhibited by human languages, can only occur just above the critical error threshold $q_a$ where the bifurcation of asymmetric and symmetric states begins. Additional empirical results from various grammar networks – data shown elsewhere (Lee, 2006; Lee *et al.*, 2007) – indicates that this threshold is highest for the complete grammar network.

Figure 5 shows the effect of graph density on the level of coherence. We plotted (a) grammar diversity, (b) dominant grammar frequency, and (c) minimum grammar frequency for a ring lattice with $n = 200$ and $a = 0.5$ given a fixed learning fidelity. Note that the learning fidelity we used ($q = 0.8$) is smaller than $q_a = 0.829$ for a complete graph. The grammar diversity and dominant frequency change in a non-linear fashion in response to network density.

When a grammar network forms a complete graph ($k = 100$), the population is in the symmetric state ($n_e = 200$) as expected. When the density of a ring lattice is sufficiently high ($k \geq 83$), the population is also in the symmetric state, and both $x_{\max}$ and $x_{\min}$ are equal to $1/n = 5 \times 10^{-3}$ as expected.

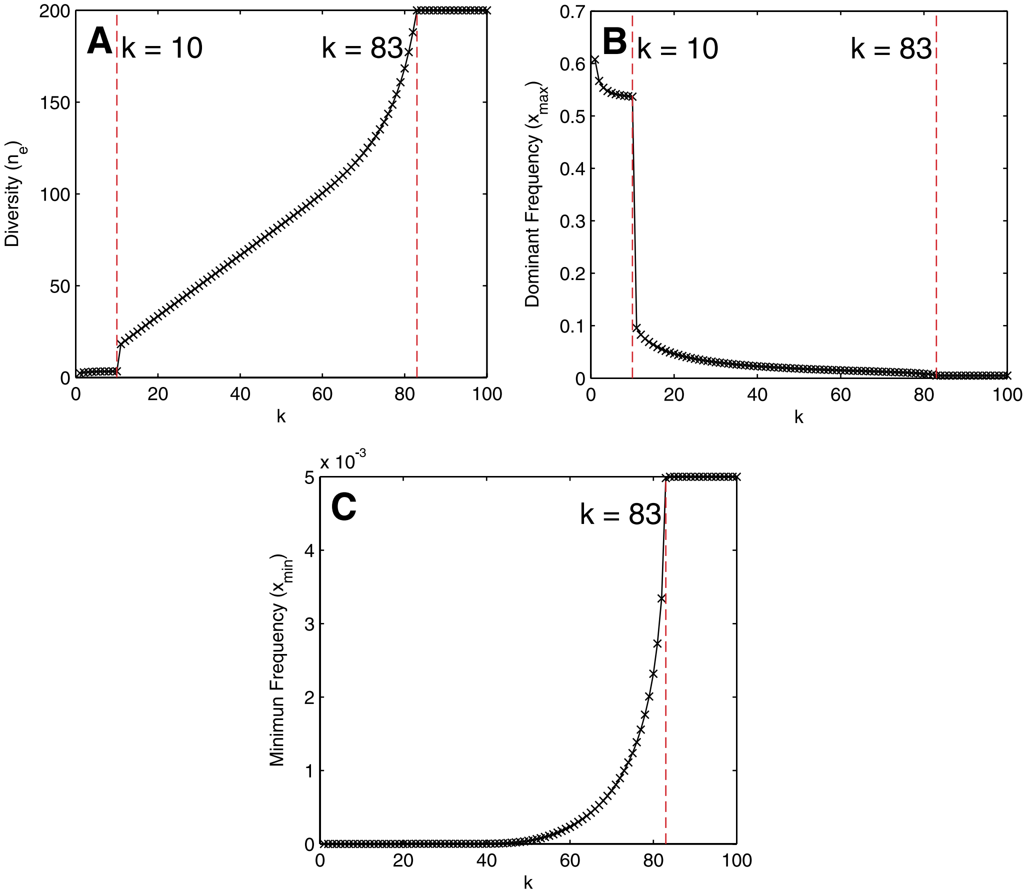

**Fig. 5.** Non-linear behaviour of language convergence in response to network density. Various measures such as (A) grammar diversity $n_e$, (B) dominant frequency $x_{\mathrm{max}}$, and (C) minimum frequency $x_{\mathrm{min}}$ are plotted for varying $k = [1; 100]$ given a ring lattice with $n = 200$, $a = 0.5$, and a learning fidelity $q = 0.8$.

As the density of the graph decreases, grammar diversity $n_e$ decreases, reflecting an increased level of linguistic coherence, indicating a weak cohesion state. When the network density is sufficiently small ($k \leq 10$), the grammar diversity is close to 1 and the dominant frequency is over 0.5, which is indicative of strong cohesion.

## DISCUSSION

In a situation where a set of grammars form a less-than-complete graph, we showed that the dynamics of the evolutionary system defined by equation (1) are characterized by three, rather than two, possible equilibrium states: (1) the symmetric state ($q \leq q_s$) where $x_i = 1/n$ for all $I$; (2) a weak cohesion state where the symmetry in grammar frequencies breaks and the distribution of grammar frequencies forms roughly a Gaussian shape centred around the most frequent grammar; and (3) a strong cohesion state ($q \geq q_a$) where a single predominant grammar emerges. We demonstrated this through an example of a ring lattice grammar network where a grammar shares a uniform similarity only among grammars that are located in less than $k$ vertices.

For a grammar space defined by a mid- to low-density regular ring lattice, a weak cohesion phase of equilibria can be identified at learning fidelities between $q_s$ and $q_a$. This region is below the error threshold for a complete graph where no cohesion or evolution can take place. The existence of a weak cohesion phase is dependent on structure in grammar space allowing the formation of a quasi-species of related grammars. While the learning fidelity in this region is too low to allow a single grammar to faithfully reproduce itself well enough to maintain a higher frequency than other grammars, the effective learning fidelity of a quasi-species as a whole is sufficiently high for the collection of grammars within it to dominate.

The existence of the weak cohesion phase suggests that a group of related grammars can emerge with lower learning fidelities than is required to establish a single consensus. Weak cohesion is also characterized by a large amount of standing heritable variation within the population, which is particularly intriguing from the evolutionary perspective.

Our results varying $k$ indicate that for regular ring lattice graphs, as the density of the grammar network decreases the level of equilibrium linguistic convergence increases for a fixed learning fidelity. We expect that this result holds generally for structured graphs, although the mean path length or clustering coefficient may be more indicative of metrics of structure in grammar space.

In technological applications where agents learn from each other and it is desirable for the overall system to converge, these results may provide a guide to designing properties of the language or state representation depending on the degree of convergence desired. If it is sufficient that agents of the system just mostly agree (i.e. converge to close variants of a dominant grammar), then a structured state space may provide a way to achieve faster convergence at higher mutation values. However, if absolute convergence is required, the state space must be designed such that minor variants are strongly selected against, producing a sharp fitness peak. This constraint also implies that a critical mutation/learning fidelity threshold exists.

The regular ring lattice is only an example of many different network structures. However, the qualitatively different dynamics with respect to the network density suggests that convergence to a set of closely related languages is significantly easier for a structured grammar space, in contrast to conclusions based only on the fully connected model.

The linguistic system described here is studied not only as an aspect of biological organisms undergoing evolution but also as an evolutionary system in itself. Situations such as gene flow among sub-populations required for speciation or cohesion into a single population are similar to the problems explored here.

## ACKNOWLEDGEMENTS

We thank Dr. Gregory M. Kobele for insightful linguistic interpretation of our model. Special thanks to Prof. Emilio Frazzoli, Prof. Jeff S. Shamma, Prof. Edward P. Stabler, and to Prof. Lawrence Slobodkin. This work was supported by MURI award administered by the US Airforce No. F49620-01-1-0361, NSF (CNS 0513058, EF-0410438), and the UCLA Center for Embedded Network Sensors. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the sponsoring agencies.

## REFERENCES

Abramowitz, M. and Stegun, C.A., eds. 1972. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 9th edn. New York: Dover.

Albert, R. and Barabasi, A.-L. (2002). Statistical mechanics of complex networks. *Rev. Modern Phys.*, **74**: 47.

Ali, A., Collier, T., Girod, L., Yao, K., Taylor, C.E. and Blumstein, D.T. (2007). An empirical study of acoustic source localization. In *IPSN '07: Proceedings of the Sixth International Conference on Information Processing in Sensor Networks*. Cambridge, MA: ACM Press (DOI: 10.1145/1236360.1236367).

Arita, T. and Taylor, C.E. (1996). A simple model for the evolution of communication. In *Proceedings of the Fifth Annual Conference on Evolutionary Programming* (L. Fogel, P.J. Angeline and T. Bäck, eds.), pp. 405–410. Cambridge, MA: MIT Press.

Collier, T.C. and Taylor, C.E. (2004). Self-organization in sensor networks. *J. Parallel Distrib. Comput.*, **64**: 866–873.

Crow, J.F. and Kimura, M. (1970). *An Introduction to Population Genetics Theory*. New York: Harper & Row.

Eigen, M. and Schuster, P. (1979). *The Hypercycle: A Principle of Natural Self-organization*. Berlin: Springer.

Eigen, M., McCaskill, J. and Schuster, P. (1989). The molecular quasi-species. *Adv. Chem. Phys.*, **75**: 149–263.

Escobar, I., Vilches, E., Vallejo, E.E., Cody, M.L. and Taylor, C.E. (2007). Self-organizing acoustic categories in sensor arrays. In *Advances in Artificial Life*, pp. 1161–1170. Berlin: Springer.

Fontana, W. and Schuster, P. (1987). A computer model of evolutionary optimization. *Biophys. Chem.*, **26**: 123–147.

Hashimoto, T. and Ikegami, T. (1995). Evaluation of symbolic grammar systems. In *Advances in Artificial Life* (F. Moran, A. Moreno, J.J. Merelo and P. Chacon, eds.), pp. 812–823. Berlin: Springer.

Hashimoto, T. and Ikegami, T. (1996). Emergence of net-grammar in communicating agents. *BioSystems*, **38**: 1–14.

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

Kobele, G.M., Riggle, J., Brooks, R., Friedlander, D., Taylor, C.E. and Stabler, E. (2004). Induction of prototypes in a robotic setting using local search MDL. In *Proceedings of the Ninth International Symposium on Artificial Life and Robotics* (M. Sugisaka and H. Tanaka, eds.), pp. 482–485. Tokyo: Springer.

Komarova, N.L. and Nowak, M.A. (2001). The evolutionary dynamics of the lexical matrix. *Bull. Math. Biol.*, **63**: 451–485.

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

Lee, Y. (2006). *The Effect of Structure in Population and Grammar Space on Language Divergence*. Los Angeles, CA: University of California.

Lee, Y., Riggle, J., Collier, T.C., Stabler, E.P. and Taylor, C.E. (2004). Adaptive communication among collaborative agents: preliminary results with symbol grouding. *Artificial Life and Robotics*, **8**: 127–132.

Lee, Y., Collier, T.C., Kobele, G.M., Stabler, E.P. and Taylor, C.E. (2005). Grammar structure and the dynamics of language evolution. In *Proceedings of Advances in Artificial Life: 8th European Conference, ECAL 2005*, pp. 624–633. Berlin: Springer.

Lee, Y., Collier, T.C., Taylor, C.E. and Stabler, E.P. (2007). Cohesion of languages in grammar networks. In *Cooperative Control of Distributed Multi-Agent Systems* (J.S. Shamma, ed.), pp. 359–375. New York: Wiley.

Marocco, D. and Nolfi, S. (2006). Emergence of communication in teams of embodied and situated agents. In *The Evolution of Language: Proceedings of the 6th International Conference on the Evolution of Language* (A. Cangelosi, A.D.M. Smith and K. Smith), pp. 198–205. Hackensack, NJ: World Scientific.

Marocco, D., Cangelosi, A. and Nolfi, S. (2003). The emergence of communication in evolutionary robots. *Phil. Trans. R. Soc. Lond. A*, **361**: 2397–2421.

Matsen, F.A. and Nowak, M.A. (2004). Win-stay, lose-shift in language learning from peers. *Proc. Natl. Acad. Sci. USA*, **101**: 18053–18057.

Niyogi, P. (2006). *The Computational Nature of Language Learning and Evolution*. Cambridge, MA: MIT Press.

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

Olfati-Saber, R. (2005). Ultrafast consensus in small-world networks. In *Proceedings of the American Control Conference*, pp. 2371–2378. New York: IEEE Press.

Olfati-Saber, R. and Murray, R.M. (2004). Consensus problems in networks of agents with switching topology and time-delays. *IEEE Trans. Automatic Control*, **49**: 1520–1533.

Rissanen, J. and Ristad, E. (1994). Language acquistion in the MDL framework. In *Language Computations* (E.S. Ristad, ed.), pp. 149–166. Philadelphia, PA: American Mathematical Society.

Slobodkin, L.B. (1992). *Simplicity and Complexity in Games of the Intellect*. Cambridge, MA: Harvard University Press.

Stabler, E.P., Collier, T.C., Kobele, G.M., Lee, Y., Lin, Y., Riggle, J. *et al.* (2003). The learning and emergence of mildly context sensitive languages. In *Advances in Artificial Life: Proceedings of the 7th European Conference, ECAL 2003* (W. Banzhaf, T. Christaller, P. Dittrich, J.T. Kim and J. Ziegler, eds.), pp. 525–534. Berlin: Springer.

Steels, L. (1996). Self-organizing vocabularies. In *Artificial Life V* (C. Langton and T. Shimohara, eds.), pp. 179–184. Nara, Japan: MIT Press.

Steels, L. (2001). Language games for autonomous robots. *IEEE Intelligent Systems*, **16**: 16–22.

Steels, L. (2006). Semiotic dynamics for embodied agents. *IEEE Intelligent Systems*, **21** (3): 32–38.