PREFINAL DRAFT: Smith, K., Brighton, H., and Kirby, S. (2003). Complex systems in language evolution: the cultural emergence of compositional structure. Advances in Complex Systems, 6(4):537–558.

Complex Systems in Language Evolution: the cultural emergence of compositional structure

Kenny Smith, Henry Brighton and Simon Kirby

Theoretical and Applied Linguistics,
School of Philosophy, Psychology and Language Sciences,
University of Edinburgh,
Adam Ferguson Building,
40 George Square,
Edinburgh EH8 9LL, UK
kenny@ling.ed.ac.uk

Abstract

Language arises from the interaction of three complex adaptive systems — biological evolution, learning, and culture. We focus here on cultural evolution, and present an Iterated Learning Model of the emergence of compositionality, a fundamental structural property of language. Our main result is to show that the poverty of the stimulus available to language learners leads to a pressure for linguistic structure. When there is a bottleneck on cultural transmission, only a language which is generalisable from sparse input data is stable. Language itself evolves on a cultural time-scale, and compositionality is language's adaptation to stimulus poverty.

1 Introduction

Human language is at the nexus of several complex adaptive systems [11]. But what are these systems, and how did they interact to deliver up language, unique among the communication systems of the natural world? In what we will call the *standard adaptationist model*, language is seen primarily as a biological trait. Language can then be explained in terms of the interaction between biological evolution of the human "language instinct" [22] and individual learning of language.

The standard adaptationist model is based on the Chomskyan paradigm from linguistics, which focuses on the innate linguistic knowledge of the speaker. However, we argue that this de-emphasis of learning and cultural transmission obscures an important dynamic in language evolution. Language itself functions as a complex adaptive system, and the historical evolution of language interacts with individual learning and biological evolution of the language faculty.

We believe that an understanding of language evolution will require a thorough understanding of each of these three complex adaptive systems (biological evolution, learning and culture), but also, crucially, an understanding of how they interact. In this paper we will focus on modelling the cultural evolution of compositionality, one of the fundamental structural characteristics of language. We present a computational model of the dynamics arising from the cultural transmission of linguistic structure. We show that compositional language can emerge from an initially non-compositional system by cultural processes. The poverty of the stimulus available to language learners drives the evolution of linguistic structure — language itself evolves to be learnable, and compositionality is language's adaptation to the poverty of the stimulus problem. In our concluding remarks we will broaden our focus to discuss how cultural evolution might interact with biological evolution and learning.

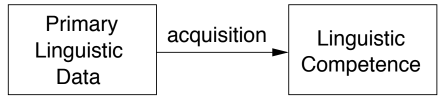

Figure 1: The Chomskyan Paradigm. The focus is on an individual's internal linguistic competence, and how this competence is acquired based on the available data. Acquisition is constrained and guided by an innate LAD and UG.

2 Complex Systems and Language Evolution

2.1 The standard adaptationist model

The standard adaptationist model places the Chomskyan approach to language within an evolutionary framework. In the Chomskyan paradigm (formulated and developed by Noam Chomsky, see for example Refs. 7 and 10), which has been highly influential in modern linguistics, language is viewed as an aspect of individual psychology. The object of interest is the internal linguistic competence of the individual, and how this linguistic competence is derived from the data the individual is exposed to. External linguistic behaviour is considered to be epiphenomenal, the uninteresting consequence of the application of this linguistic competence to a set of contingent communicative situations. From this standpoint, much of the structure of language is puzzling — how do children, apparently effortlessly and with virtually universal success, arrive at a sophisticated knowledge of language from exposure to sparse and noisy data? In order to explain language acquisition in the face of this poverty of the linguistic stimulus, the Chomskyan program postulates a sophisticated, genetically-encoded language organ of the mind, consisting of a Universal Grammar (UG), which delimits the space of possible languages, and a Language Acquisition Device (LAD), which guides the formation of linguistic competence based on the observed data. Language learners are therefore viewed as detached individuals [5], as illustrated in Figure 1.

Chomsky has been notoriously reluctant to offer an account of the evolution of UG and the LAD, preferring instead to appeal to architectural and developmental constraints:

> "We know very little about what happens when 10^{10} neurons are crammed into something the size of a basketball, with further conditions imposed by the specific manner in which this system developed over time. It would be a serious error to suppose that all properties, or the interesting properties of the structures that evolved, can be 'explained' in terms of natural selection." [8, p59]

However, others have been less reticent in attempting to integrate the Chomskyan paradigm with evolutionary theory. Pinker & Bloom present the classic adaptationist account of language evolution, suggesting that "the ability to use a natural language belongs more to the study of human biology than human culture: it is a topic like echolocation in bats"[23, p707]. They argue that language is adapted for the communication of propositional structures (in the internal representational "language of thought") over a serial channel. UG and the LAD have therefore evolved to facilitate the acquisition of language which performs this function.

Pinker & Bloom's account calls upon the interaction between two complex adaptive systems to explain the language capacity and linguistic structure. The process of language acquisition, constrained by UG and guided by the LAD, determines an individual's linguistic competence. This competence then contributes to an individual's reproductive fitness, resulting in selection in favour of an innate endowment which 1) facilitates language acquisition and 2) constrains the learner to learning languages which are communicatively useful. Biological adaptation of UG and the LAD then feeds back into the language acquisition process. This interaction is illustrated in Figure 2.

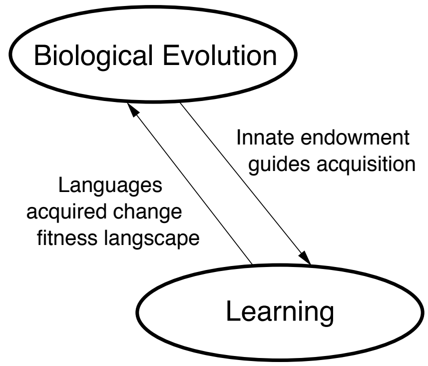

Figure 2: The standard adaptationist model. Language is a consequence of the interaction between biological evolution and learning. The innate language capacity guides language acquisition. The functionality of the acquired language then has consequences for the biological evolution of the language capacity.

2.2 Culture: a third complex system

Those working within the Chomskyan paradigm typically play down the role of learning and the cultural transmission of language. For Chomsky, learning of a language is "better understood as the growth of cognitive structures along an internally directed course under the triggering and partially shaping effect of the environment" [9, p34], while Piattelli-Palmarini has suggested that "we would gain in clarity if the *scientific* use of the term [learning] were simply discontinued"[21, p2]. This devaluation of learning and, consequently, cultural transmission arises from concerns based on poverty of the stimulus arguments. If the linguistic stimulus available to the child is too impoverished to allow language acquisition, then much of the structure of language must be prespecified, and learning and culture effectively play no role. However, focusing on the nature of this innate knowledge, to the detriment of the study of the cultural transmission of language, means that we overlook an important dynamic which can help explain some of the fundamental structural properties of language.

Following ideas developed by Hurford [15], we place an understanding of cultural evolution at the heart of our explanatory approach. An individual's linguistic competence is derived from data which is itself a consequence of the linguistic competence of other individuals. This view of language is illustrated in Figure 3.

What consequences does this view of language have for evolutionary explanations of language and the language faculty? The introduction of cultural transmission results in a third complex adaptive system, that of cultural evolution, operating on what has been dubbed a *glossogenetic* [15] time-scale, intermediate between the phylogenetic and ontogenetic time-scales. As in the standard adaptationist model, language acquisition is guided by an individual's innate endowment. The learner attempts to acquire the language of their cultural parents. Differences between the language of the parent and the child results in the cultural evolution of language itself. This cultural evolution further restricts the set of possible languages available to language learners at subsequent generations. The particular language acquired by a learner from the set of languages made available by this cultural evolution then contributes to the reproductive fitness of that individual, resulting in selection in favour of an innate endowment which 1) facilitates acquisition of those languages present in the culture and 2) constrains the learner to learning languages which are communicatively useful. Biological adaptation of UG and the LAD then feeds back into the language acquisition process. This interaction is illustrated in Figure 4.

The structure of language is dependent on the interaction between these three complex systems, and a full understanding of language evolution will require a treatment of all three adaptive

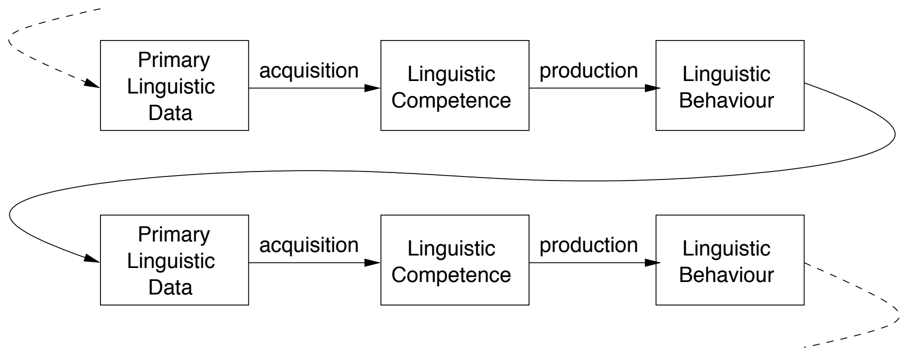

Figure 3: Language as a cultural phenomenon. As in the Chomskyan paradigm, illustrated in Figure 1, acquisition based on available data leads to linguistic competence. Importantly, however, this competence in turn leads to linguistic behaviour, which becomes the linguistic data for the next generation of language learners.

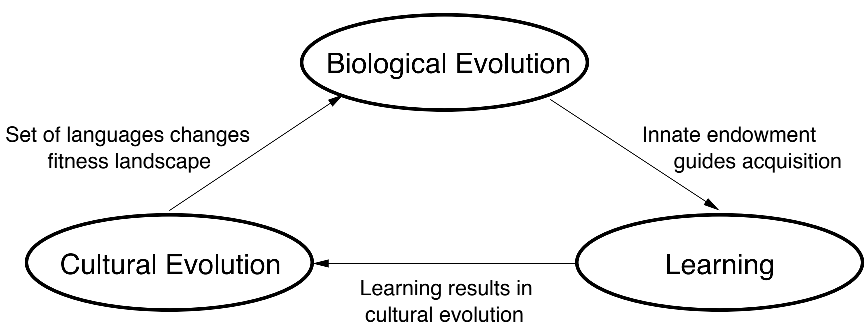

Figure 4: Adding cultural evolution. Language is now a consequence of the interaction between biological evolution, learning and cultural evolution. The innate language capacity guides language acquisition. The cultural transmission of language leads to cultural evolution, which then has consequences for the biological evolution of the language capacity.

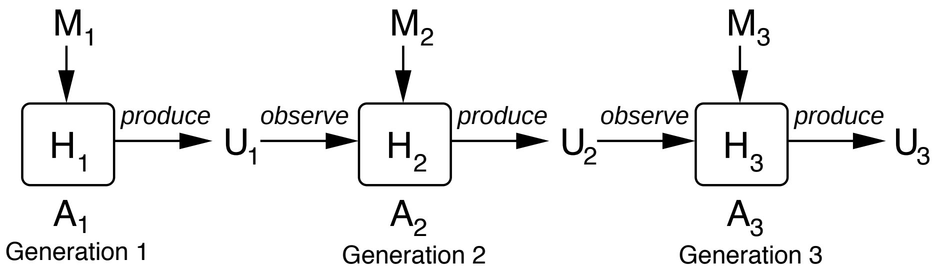

Figure 5: The ILM. In the simplest case, the ith generation of the population consists of a single agent A_i who has hypothesis H_i. Agent A_i is prompted with a set of meanings M_i. For each of these meanings the agent produces an utterance using H_i. This yields a set of utterances U_i. Agent A_{i+1} observes U_i and forms a hypothesis H_{i+1} to explain the set of observed utterances, and the cycle repeats.

processes. However, it is prudent to develop an understanding of each process in isolation before attempting to formulate a complete, unified model of the evolution of language. In this paper we will develop an account of the dynamics arising from the cultural transmission of language, then draw inferences as to how this dynamic might interact with the complex systems of individual learning and biological evolution.

3 The Iterated Learning Model

The Iterated Learning Model (ILM) provides a framework for studying the cultural evolution of language on a glossogenetic time-scale [18, 4]. The ILM in its simplest form is illustrated in Figure 5. In this model H_i corresponds to the linguistic competence of individual i, whereas U_i corresponds to the linguistic behaviour of individual i and the primary linguistic data for individual $i+1$.

We focus here on the cultural evolution of language in the absence of any functional pressure for effective communication. While it has been suggested that functional considerations have an impact on language acquisition and production (for example, a preference by speaker/hearers for sentences which are easy to parse [14]), ignoring such pressures allows us to make several simplifying assumptions. We can treat the population at any given generation as consisting of a single agent. This means that we can focus fully on vertical cultural transmission, and ignore for the moment horizontal, within-generation transmission. We can also ignore inter-generational communication. However, the ILM does not rule out a focus on the communicative function of language within or between generations in a population (see, for example, Ref. 25) or the role of horizontal transmission (see Ref. 2 for an ILM where transmission is purely horizontal).

The ILM provides a powerful framework for investigating the cultural evolution of language. We have previously used the ILM to examine the emergence of word-order universals [17], the regularity–irregularity distinction [18] and recursive syntax [19]. Here we will focus on the cultural evolution of compositionality, one of the characteristic structural properties of language.

4 Modelling the Evolution of Compositionality

4.1 Compositionality

In a compositional communication system the meaning of a signal is a function of the meaning of its parts [20]. The morphosyntax of human language is highly compositional. For example, the relationship between the sentence *John walked* and its meaning is not completely arbitrary. The sentence is made up of two components: a noun (*John*) and a verb (*walked*). The verb is also made up of two components: a stem and a past-tense ending. The meaning of *John walked* is thus a function of the meaning of its parts. Compositionality, in combination with recursive syntax, allows language users to produce and comprehend an infinite range of sentences.

Compositional language can be contrasted with non-compositional, or *holistic* communication, where a signal stands for the meaning as a whole, with no subpart of the signal conveying any part of the meaning in and of itself. Animal communication is typically viewed as holistic — no subpart of an alarm call or a mating display stands for part of the meaning "there's a predator about" or "come and mate with me".

How can we explain the compositionality of language? In the standard adaptationist model, compositionality must be viewed as a consequence of a biological adaptation of the unique human language organ — natural selection has favoured an innate endowment which restricts language learners to learning compositional systems. However, we demonstrate that compositionality can arise through purely cultural processes, as a result of the adaptation of language in the face of pressure to be learnable. This lifts some of the burden of explanation from the postulated language organ — cultural processes acting on a (possibly domain-general) biological substrate result in compositional language.

4.2 An analysis of stable states

Brighton has developed a mathematical analysis of the relative stability of compositional and holistic language over cultural time [4]. Brighton considers only perfectly compositional and completely holistic language. Addressing the question of relative stability allows us to predict when we should observe linguistic structure — when compositional and holistic language are equally stable we should expect them to emerge with equal frequency over cultural time, whereas when one type of language is more stable than the other we should expect that language to emerge more frequently and persist for longer.

One of Brighton's observations is that, in an iterated learning scenario, stability of a language over cultural time relates to the expressivity of learners exposed to that language. Consider the problem faced by a learner attempting to learn a holistic language. Given the lack of structure in the holistic language, the best strategy for the learner is simply to memorise meaning-signal pairs. The learner, when called upon to produce an utterance, will only be able to faithfully reproduce meaning-signal pairs that it itself has observed. Parts of the language which have not been observed cannot be expressed and will therefore be lost or will change — holistic language is only stable when the learner observes, and is therefore able to express, the complete language of the previous generation.

In contrast, the structure of a compositional language means that learners can acquire and express the complete language based on observation of a subset of that language. Consider a learner presented with a perfectly compositional language. In such a language each element of meaning will map onto a particular part of signal (for example, an affix or a word). The best strategy here is to memorise the association between elements of meaning and parts of signal. When called upon to produce an utterance for a given meaning, an individual is not restricted to reproducing meaning-signal pairs it itself has observed. A meaning will be expressible if every element of that meaning has been observed paired with its associated linguistic unit. A learner of a compositional language can therefore generalise from observed examples to express parts of the language that it has not actually observed — incomplete exposure to the target language does not result in a shortfall in expressivity, and therefore the language will remain stable.

Brighton's key result is to show that the stability advantage of compositional language over holistic language is at a maximum when there is a *bottleneck* on cultural transmission. The transmission bottleneck occurs when learners only observe a subset of the language of the previous generation. This is one aspect of the poverty of the stimulus problem — the set of utterances of any human language is arbitrarily large, but a child must acquire their linguistic competence based on a finite number of sentences. The severity of the transmission bottleneck is given by the proportion of the language of the previous generation that a learner will observe.

Holistic languages cannot persist over time when the bottleneck on cultural transmission is tight — learners can only faithfully reproduce parts of the language which they have observed, and if they observe only a small subset of the language then the language will be unstable. In contrast, compositional languages are generalisable, due to their structure, and remain relatively stable even when a learner only observes a small subset of the language of the previous generation. Brighton shows that the poverty of the stimulus "problem" is actually a requirement for linguistic structure — were there no poverty of the stimulus, compositional language would have no advantage over unstructured holistic language.

4.3 A model of language dynamics

Brighton's result is a fundamental one. However, by considering only perfectly compositional or completely holistic languages, Brighton is restricted to examining the Lyapounov stable states, places in language space that, if we start near, we stay near [12]. The model cannot explain the dynamics that occur when we move away from the extremes of compositionality, although insights taken from the model prove relevant to understanding the behaviour of dynamic models.

What happens to languages of intermediate compositionality during cultural transmission? Can compositional language emerge from initially holistic language, through a process of cultural evolution? We can investigate these question using a multi-agent computational implementation of the ILM. A simple model of language as a mapping between meanings and signals is given in Section 4.3.1. A neural network model of a linguistic agent capable of learning and producing such languages, based on a simple model designed to investigate the cultural evolution of vocabulary systems [25], is outlined in Section 4.3.2. This agent is inserted into the ILM, along with a model of environments (Section 4.3.3), allowing us to model the dynamics arising from the cultural transmission of language.

4.3.1 The language model

We treat language as a mapping between meanings and signals. A compositional language is a mapping which preserves neighbourhood relationships — neighbouring meanings will share structure, and that shared structure in meaning space will map to shared structure in signal space. A holistic language is one which does not preserve such relationships — as the structure of signals does not reflect the structure of the underlying meaning, shared structure in meaning space will not necessarily result in shared signal structure.

In order to model such systems we need representations of meanings and signals. Meanings are represented as points in an F-dimensional space where each dimension has V discrete values, and signals are represented as strings of characters of length 1 to l_{max}, where the characters are drawn from some alphabet Σ. More formally, the meaning space $\mathcal{M}$ and signal-space $\mathcal{S}$ are given by:

$$\mathcal{M} = \{(f_1\ f_2 \dots f_F) : 1 \leq f_i \leq V \text{ and } 1 \leq i \leq F\} \tag{1}$$

$$\mathcal{S} = \{w_1 w_2 \dots w_l : w_i \in \Sigma \text{ and } 1 \leq l \leq l_{max}\} \tag{2}$$

Utterances, the units of observable behaviour that individuals acquire their competence from, are considered to be meaning-signal pairs $\langle m, s\rangle$, where $m \in \mathcal{M}$ and $s \in \mathcal{S}$. We therefore assume that learners are able to deduce the communicative intentions of others during language acquisition. This is obviously an oversimplification — if the meaning of every signal was self-evident then the signal itself would serve little purpose. However, we can make several points in defence of this idealisation. Firstly, children do seem to have various strategies for deducing the meaning underlying an observed signal. Central to these abilities is the capacity to establish joint attention and perform intentional inference [1, 3]. Secondly, computational simulations show that linguistic structure can be preserved if this idealisation is weakened, so that learners observe only partially-specified meanings in conjunction with signals [16]. Finally, a strand of research parallel and complementary to our own abandons this idealisation completely (see, for example, Refs. 24, 26, 27, 28). This work shows that shared linguistic structure can still emerge in the absence of explicit meaning transfer during learning.

We make a fundamental distinction between meanings and signals. An alternative approach is presented by Hashimoto [13], who also treats language as a dynamical system but makes no distinction between the meanings of words and the words themselves. From this standpoint, the meaning of a signal is defined in terms of the relationship between that signal and other signals — the mesh of word-word associations constrains and guides the interpretation of signals. Language lies somewhere between these two extremes.

4.3.2 A network model of a linguistic agent

Representation Agents are modelled using networks consisting of two sets of nodes $\mathcal{N}_M$ and $\mathcal{N}_S$ and a set of bidirectional connections $\mathcal{W}$ connecting every node in $\mathcal{N}_M$ with every node in $\mathcal{N}_S$. Nodes in $\mathcal{N}_M$ represent meanings and partial specifications of meanings, while nodes in $\mathcal{N}_S$ represent partial and complete specifications of signals.

As summarised above, each meaning is a vector in F-dimensional space where each dimension has V values. *Components* of meanings are (possibly partially specified) vectors, with each feature of the component either having the same value as the meaning, or a wildcard. More formally, if c_m is a component of meaning m, then the value of the jth feature of c_m is:

$$c_m[j] = \begin{cases} m[j] & \text{for specified features} \\ * & \text{for unspecified features} \end{cases} \tag{3}$$

where $*$ represents a wildcard. Similarly, components of signals of length l are (possibly partially specified) strings of length l. We impose the additional constraint that a component must have a minimum of one specified position. For example, the components of the meaning represented by the vector (1 2) are (1 2), (1 $*$) and ($*$ 2), but not (1 3) (value of feature 2 doesn't match) or ($*$ $*$) (no specified features). Similarly, the components of the signal represented by the string *bd* are *bd*, *b*$*$ and $*$*d*, but not *e*$*$ (first character doesn't match), $**$ (no specified characters) or *a* (not of correct length).

Each node in $\mathcal{N}_M$ represents a component of a meaning, and there is a single node in $\mathcal{N}_M$ for each component of every possible meaning. Similarly, each node in $\mathcal{N}_S$ represents a component of a signal and there is a single node in $\mathcal{N}_S$ for each component of every possible signal.

Learning During a learning event, a learner observes a meaning-signal pair $\langle m, s\rangle$. The activations of the nodes corresponding to all possible components of m and all possible components of s are set to 1. The activations of all other nodes are set to 0. The weights of the connections in $\mathcal{W}$ are adjusted according to the weight-update rule:

$$\Delta W_{xy} = \begin{cases} +1 & \text{if } a_x = a_y = 1 \\ -1 & \text{if } a_x \neq a_y \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

where W_{xy} gives the weight of the connection between nodes x and y and a_x gives the activation of node x. The learning procedure is illustrated in Figure 6.

Production An *analysis* of a meaning or signal is an ordered set of components which fully specifies that meaning or signal. More formally, an analysis of a meaning m is a set of N components $\{c_m^1, c_m^2, \ldots c_m^N\}$ that satisfies two conditions:

1. If $c_m^i[j] = *$, $c_m^k[j] \neq *$ for some choice of $k \neq i$
2. If $c_m^i[j] \neq *$, $c_m^k[j] \neq c_m^i[j]$ for any choice of $k \neq i$

The first condition states that an analysis may not consist of a set of components which all leave a particular feature unspecified — an analysis fully specifies a meaning. The second states that an analysis may not consist of a set of components where more than one component specifies the value of a particular feature — analyses do not contain redundant components. Valid analyses of signals are similarly defined.

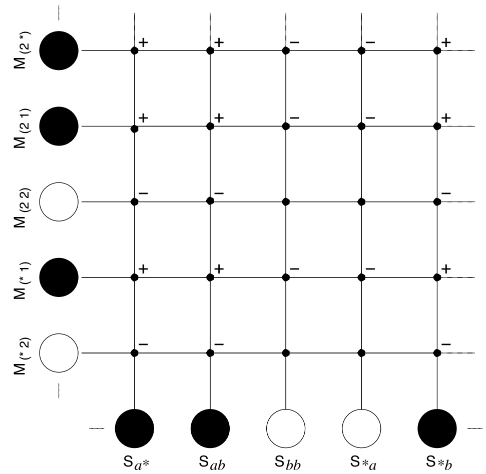

Figure 6: Storage of the meaning-signal pair $\langle(2\ 1), ab\rangle$. Nodes with an activation of 1 are represented by large filled circles, and are labelled with the component they represent. For example, $M_{(2\ *)}$ is the node which represents the meaning component $(2\ *)$. Small filled circles represent weighted connections. During the learning process, nodes representing components of $(2\ 1)$ and ab have their activations set to 1. Connection weights are then either incremented $(+)$, decremented $(-)$ or left unchanged.

During the process of producing utterances, agents are prompted with a meaning and required to produce a meaning-signal pair. Production proceeds via a winner-take-all process. In order to retrieve a signal $s_i \in \mathcal{S}$ based on an input meaning $m_i \in \mathcal{M}$ every possible signal $s_j \in \mathcal{S}$ is evaluated with respect to m_i. For each of these possible meaning-signal pairs $\langle m_i, s_j\rangle$, every possible analysis of m_i is evaluated with respect to every possible analysis of s_j. The evaluation of a meaning analysis-signal analysis pair yields a score g. The meaning-signal pair which yields the analysis pair with the highest g is returned as the network's production for the given meaning. The score for a meaning analysis (which consists of a set of meaning components) paired with a signal analysis (a set of signal components) is given by:

$$g\left(\left\{c_m^1, c_m^2 \ldots c_m^N\right\}, \left\{c_s^1, c_s^2 \ldots c_s^N\right\}\right) = \sum_{i=1}^{N} \omega\left(c_m^i\right) \cdot W_{c_m^i, c_s^i} \tag{5}$$

where N is the number of components in the analysis of meaning and signal, $w_{c_m^i, c_s^i}$ gives the weight of the connection between the nodes representing the ith component of the meaning analysis and the ith component of the signal analysis and $\omega(x)$ is a weighting function which gives the non-wildcard proportion of x. The production process is illustrated in Figure 7.

4.3.3 Environments

The world, which provides communicatively relevant situations for agents in our model, consists of a set of objects, where each object is labelled with a meaning drawn from $\mathcal{M}$. We will refer to such a set of labelled objects as an *environment*. The number of objects in the environment gives the *density* of that environment — environments with few objects will be termed low-density, whereas environments with a large number of objects will be termed high-density. When meanings are assigned to objects at random we will say the environment is *unstructured*. When meanings are assigned to objects in such a way as to minimise the average inter-meaning Hamming distance we will say the environment is *structured*. Sample low- and high-density environments are shown in Figure 8.

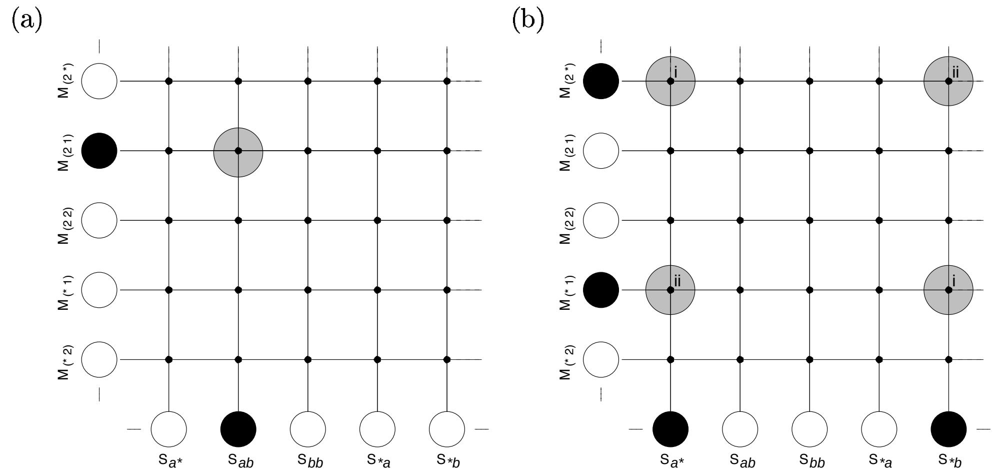

Figure 7: Retrieval of three possible analyses of $\langle(2\ 1)\,, ab\rangle$. The relevant connection weights are highlighted in grey. (a) g for the one-component analysis $\langle\{(2\ 1)\}\,, \{ab\}\rangle$ depends on the weight of the connection between the nodes representing the components (2 1) and ab. (b) g for the two-component analysis $\langle\{(2\ *)\,, (*\ 1)\}\,, \{a*, *b\}\rangle$ depends on the weighted sum of two connections, marked as i. The g for the alternative two-component analysis $\langle\{(2\ *)\,, (*\ 1)\}\,, \{*b, a*\}\rangle$ is given by the weighted sum of the two connections marked ii.

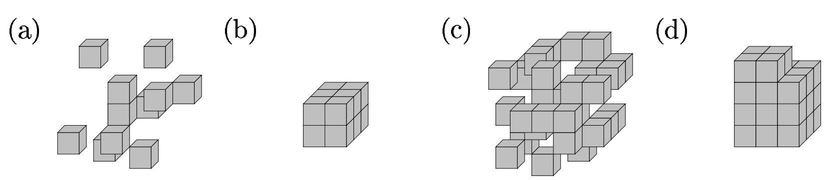

Figure 8: We will present results for the case where $F = 3$ and $V = 5$. This defines a three-dimensional meaning space. We highlight the meanings selected from that space with grey. (a) is a low-density, unstructured environment. (b) is a low-density, structured environment. (c) and (d) are unstructured and structured high-density environments.

4.3.4 Measuring compositionality

As discussed above, a compositional mapping preserves neighbourhood relations when mapping between meanings and signals, whereas a holistic mapping does not, unless by chance. Our measure of compositionality, c, captures this and is calculated based on the set of meaning-signal pairs in an agent's language. c is the Pearson's Product-Moment correlation coefficient of the pairwise distances between pairs of meanings and the distance between the corresponding pairs of signals. We use the Hamming and Levenstein (string edit) distance measures to quantify inter-meaning and inter-signal distances respectively. c ranges between -1 and 1. A perfectly compositional language will have a c of 1, whereas $c \approx 0$ for holistic languages.

5 Results

The network model of a linguistic agent outlined above is plugged into the ILM framework described in Section 3. We will vary three key parameters — the presence or absence of a bottleneck on cultural transmission, and the density and structure of the environment.

5.1 No bottleneck on cultural transmission

First, runs of the ILM were carried out in the absence of a bottleneck on cultural transmission — each learner is presented with the complete language of the agent at the previous generation. The initial agents have all their connections weights set to 0, and therefore produced every meaning-signal pair with equal probability. Subsequent agents have connection weights of 0 prior to learning. Runs were allowed to progress until a stable state was reached, where agent A_i and A_{i+1} produced identical languages. At this point, in the absence of a bottleneck, further language change is impossible.

Figures 9 and 10 plot compositionality by frequency for the initial and final, stable languages for low-density and high-density environments. These results are based on 1000 independent runs of the ILM for each environment.

Three results are apparent from these Figures:

1. Highly compositional systems are infrequent.

2. Compositional systems only occur when the environment is low-density.

3. Highly compositional systems only occur when the environment is structured.

The initial languages in all environments tend to be holistic. Previous results [4] suggest that, with no bottleneck on cultural transmission, such systems will be highly stable. This seems to be the case for the majority of runs reported here, particularly in the high-density environments. The emergence of partially or highly compositional systems in the low-density environments then seems somewhat surprising.

Individual simulation runs can be split into three groups – those where the final languages have the same level of compositionality as the initial languages ($c_{initial} = c_{final}$), those where the final compositionality is different from the compositionality of the initial language but not high ($c_{initial} \neq c_{final}$, $c_{final} < 0.9$), and those where the final systems are highly compositional ($c_{initial} \neq c_{final}$, $c_{final} \geq 0.9$). Table 1 gives the mean and standard deviations of the compositionality of the initial languages from the simulation runs, organised into these three groups.

As can be seen from the second column of the Table, runs in all environments have a mean value of $c_{initial}$ of approximately 0. However, these initial values are much more tightly distributed around the mean in the high-density environments. The third column gives the mean $c_{initial}$ for simulation runs where $c_{initial} = c_{final}$. These values are somewhat lower than the overall mean, and are lower than the mean $c_{initial}$ for simulation runs which move away from initial value. The mean $c_{initial}$ for simulation runs which converge on highly compositional languages is higher still.

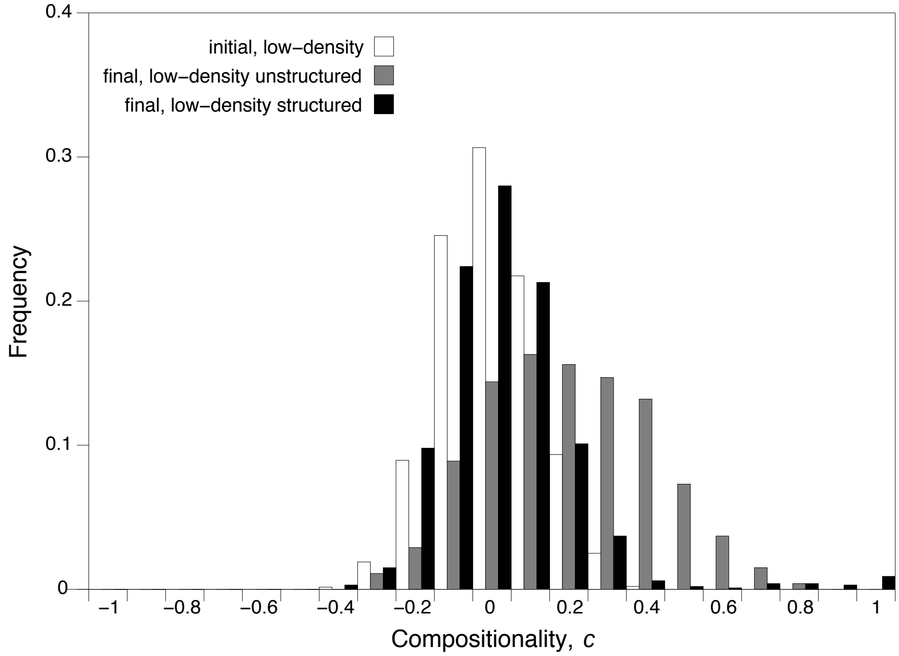

Figure 9: The relative frequency of initial and final systems of varying degrees of compositionality, where there is no bottleneck on cultural transmission. The results shown here are for the low-density environments given in Figure 8. The initial languages are largely holistic. Partially compositional languages do emerge, but highly compositional languages are infrequent.

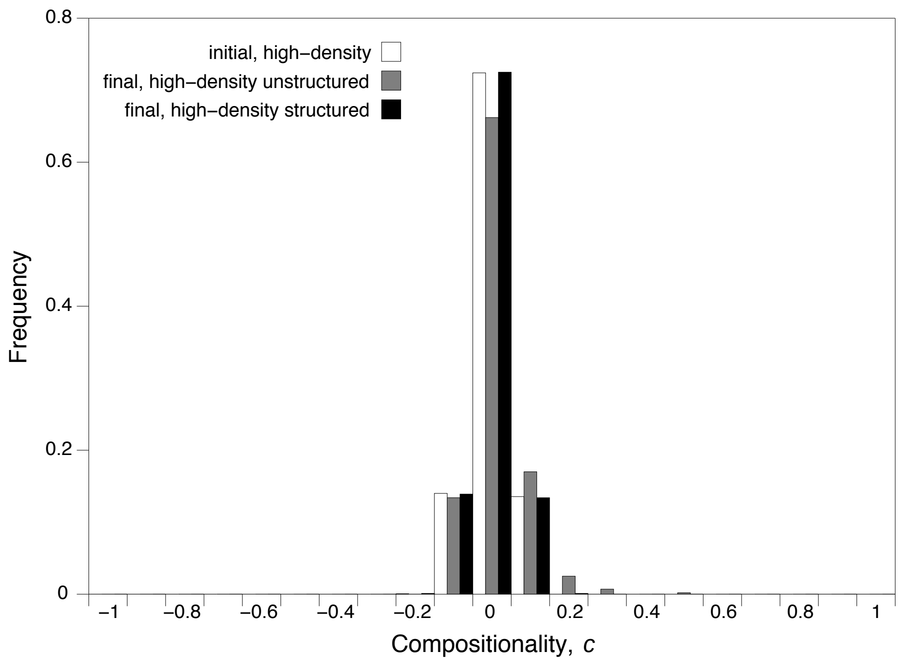

Figure 10: The relative frequency of initial and final systems of varying degrees of compositionality, where there is no bottleneck on cultural transmission. The results shown here are for the high-density environments given in Figure 8. Both the initial and final languages are holistic.

Environment	Initial Compositionality			
	all	$c_{initial} = c_{final}$	$c_{initial} \neq c_{final}$, $c_{final} < 0.9$	$c_{initial} \neq c_{final}$, $c_{final} \geq 0.9$
ld, u	$\mu = 0.0013$, $\sigma = 0.1246$	$\mu = -0.0512$	$\mu = 0.0160$	NA
ld, s	$\mu = -0.0029$, $\sigma = 0.1246$	$\mu = -0.0136$	$\mu = 0.0536$	$\mu = 0.1603$
hd, u	$\mu = -0.0004$, $\sigma = 0.0470$	$\mu = -0.0047$	$\mu = 0.0209$	NA
hd, s	$\mu = -0.0011$, $\sigma = 0.0457$	$\mu = -0.0017$	$\mu = 0.0306$	NA

Table 1: Sensitivity to initial conditions. Environments are specified by a density — low-density (ld) or high-density (hd), and a degree of structure — unstructured (u) or structured (s). Simulation runs can be split into three groups, according to the values of $c_{initial}$ and c_{final}. The table gives the means (μ) of the initial language of the simulation runs, broken down by group. Standard deviations (σ) are given once as σ for each subgroup is approximately the same as σ for the runs in that environment as a whole.

Also note that the mean value of $c_{initial}$ is lower for unstructured environments than for structured environments for the case where $c_{initial} \neq c_{final}$.

These results suggest that, in the absence of a bottleneck on cultural transmission, there is a degree of sensitivity to the compositionality of the initial, random language. Where this initial system exhibits compositional tendencies, yielding $c_{initial}$ above the mean, there is an increased likelihood of the system moving, over iterated learning events, towards more compositional regions of language space. The compositional tendencies of the initial system spread to other parts of the system over time, resulting in an increase in compositionality. For the high-density environments, highly compositional systems do not emerge due to the fact that the initial systems tend to be clustered more tightly around the non-compositional mean. When the environment contains few meanings the initial system may, by chance, exhibit some compositional tendencies. However, when the environment contains a large number of meanings such tendencies are likely to be drowned out by the majority non-compositional mapping.

Why does environment structure impact on the compositionality of languages in the low-density environments? This is related to the previous point. In the low-density environments, as discussed above, compositional tendencies in the initial system spread, over time, to other parts of the system. In structured environments, distinct meanings tend to have feature values in common with a large number of other meanings. In unstructured environments distinct meanings have feature values in common with few other meanings. If the initial random system has a tendency to express a given feature value with a certain substring then this can spread to cover all meanings involving that feature value — the system becomes consistent with respect to that feature value, which can have knock-on consequences for other values at that feature and other features. In structured environments the potential for spread of the substring associated with a particular feature value is wider than is the case in unstructured environments, given that more meanings will share that feature value. Any initial compositional tendency will therefore spread more widely in structured environments, with more possible follow-on consequences, resulting in the more frequent emergence of highly compositional languages.

However, while shared feature values allow the possibility of the spread of compositionality, they also inhibit it — in a structured environment, any compositional tendency in the initial random mapping has to cover a large number of meanings which share feature values. If only some of these meanings share a character for a particular feature value, then the other meanings, which do not share the character, are likely to outweigh the slight compositional tendency. In contrast, in unstructured environments fewer meanings share feature values, therefore the initial random system has to be less 'lucky' in the assignment of characters to feature values. This is reflected in the fact that the mean $c_{initial}$ has to be higher in structured environments before c_{final} moves away from $c_{initial}$, and also in the fact that the average c_{final} in unstructured environments is higher (see Figure 9). In structured environments, the initial compositional tendency has to be strong to escape the attraction of the overall non-compositional mapping, but once this attraction has

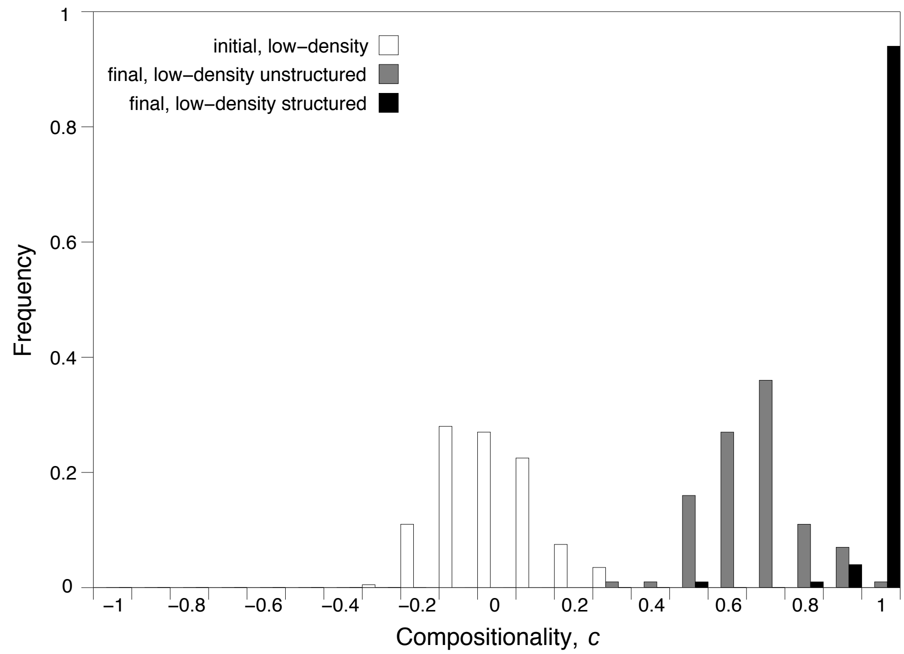

Figure 11: The relative frequency of initial and final systems of varying degrees of compositionality, where there is a bottleneck on cultural transmission. The results shown here are for the low-density environments. The initial languages are largely holistic. Partially and highly compositional languages emerge with high frequency, and are highly compositional languages are most frequent when the environment is structured.

been escaped highly compositional systems can emerge. In contrast, in unstructured environments the attraction of the initial non-compositional mapping is weaker, due to the reduced degree of feature-value sharing, but the potential spread of compositionality is reduced.

5.2 Bottleneck on cultural transmission

Next, runs of the ILM were carried out with a bottleneck on cultural transmission — each learner is presented with a subset of the language of the agent at the previous generation. The number of utterances produced by agents was set so that language learners observed utterances for approximately 40% of the language of the previous agent.

While in the absence of a bottleneck runs were allowed to proceed until a stable state was reached, in the bottleneck condition runs were terminated after a fixed number of generations (200). The random selection of objects from the environment for which to produce utterances means that, as with any stochastic system, a highly skewed distribution of objects could lead to the loss of structure. However, the results reported here accurately reflect the behaviour of the system — allowing the runs to proceed for several hundred more generations gives a similar distribution of languages.

Figures 11 and 12 plot the compositionality by frequency of the initial and final languages for unstructured and structured low and high-density environments. These results are based on 100 independent runs of the ILM for each environment. Fewer runs are required as the transmission bottleneck reduces the sensitivity to initial conditions.

Two results are apparent from Figures 11 and 12:

1. Highly compositional systems are frequent.

2. Highly compositional systems are most frequent when the environment exhibits structure.

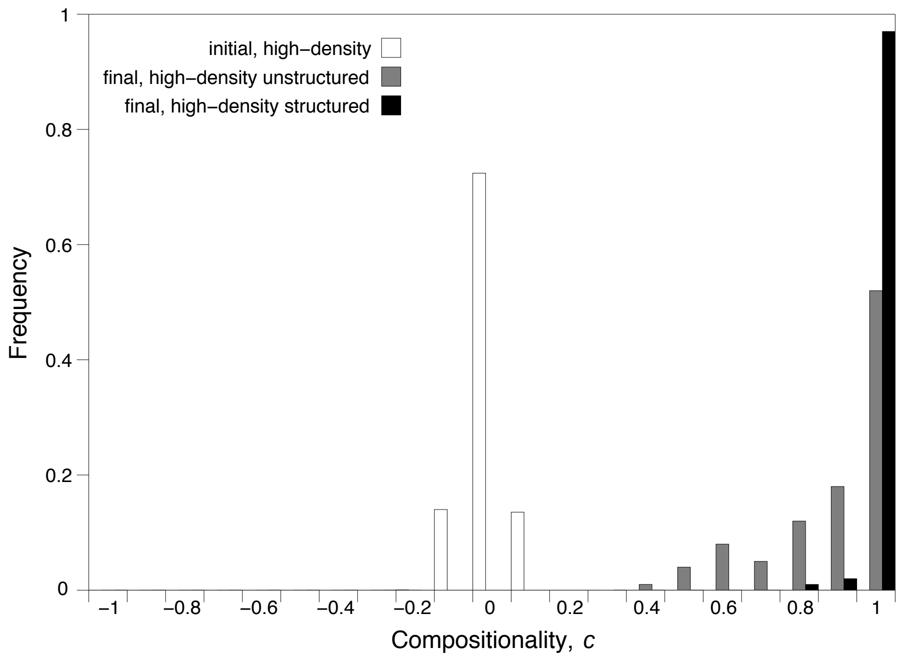

Figure 12: The relative frequency of initial and final systems of varying degrees of compositionality, where there is a bottleneck on cultural transmission. The results shown here are for the high-density environments. The initial languages are largely holistic. Highly compositional languages emerge with high frequency, and are most frequent when the environment is structured.

Brighton's mathematical model predicts that, in the presence of a bottleneck on cultural transmission, compositional language will be more stable than holistic language [4]. The results from the computational model bear this out, but also show that it is possible to move from an initially holistic system to a highly compositional system over time. Figures 13 and 14 illustrates the dynamics of the transition from holistic to compositional language. In structured environments (of low or high-density) there is a single attractor at $c = 1$. Systems reaching this point are highly stable, and any perturbation away from the attractor is quickly reversed. In contrast, in unstructured environments the attractor either occurs at a lower level of compositionality (as in the low-density unstructured environment), or the attractor occurs at $c = 1$ but is approached more slowly and has a slight repellent effect (as in the high-density unstructured environment).

Why are compositional languages so strongly preferred when there is a bottleneck on transmission? Holistic languages cannot persist in the presence of a bottleneck. The meaning-signal pairs of a holistic language have to be observed to be reproduced. When a learner only observes a subset of the holistic language of the previous generation then certain meaning-signal pairs will not be preserved — the learner, when called upon to produce, will produce some other signal for that meaning, resulting in a change in the language. In contrast, compositional languages are generalisable, due to their structure, and remain relatively stable even when the learner observes a small subset of the language of the previous generation. Over time, language adapts to the pressure to be generalisable. Eventually, particularly when the environment is structured, the language becomes highly compositional, highly generalisable and consequently highly stable.

In a structured environment the advantage of compositionality is at a maximum. As discussed above, in such environments meanings share feature values with several other meanings. A language mapping these feature values to a signal substring is highly generalisable. When the environment is unstructured, meanings share feature values with few or no other meanings. In the most extreme case, a meaning may have a value for a particular feature which no other meaning has. The signal associated with that meaning cannot then be deduced from observations of

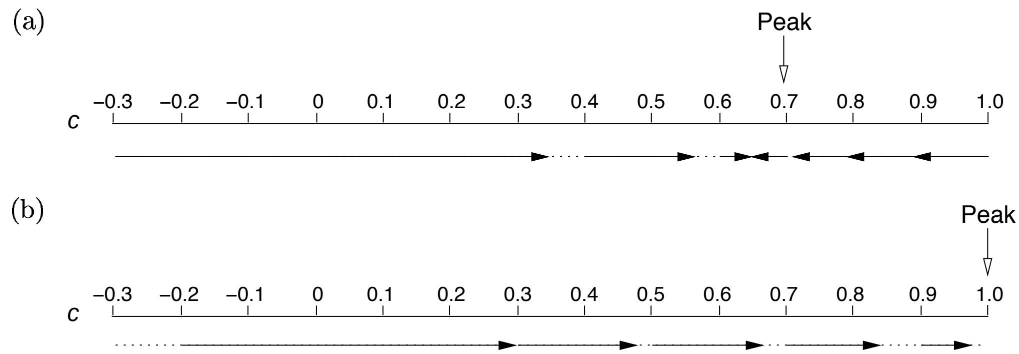

Figure 13: The dynamics of language change in low-density environments. Arrows represent the direction and magnitude of change of languages of a given level of compositionality, c. The origin of the arrow gives the compositionality of the language at time t. The direction and length of the arrow corresponds to the mean directionality and magnitude of change in compositionality for those systems at time $t+1$. (a) Dynamic in the low-density unstructured environment. There is an attractor, around $c = 0.7$, corresponding to the peak of the distribution given in Figure 11. Magnitude of change decreases as this attractor is approached. (b) Dynamics in the low-density structured environment. There is consistent movement towards the attractor at $c = 1$, corresponding to the peak in Figure 11. Again, magnitude of change decreases as this attractor is approached.

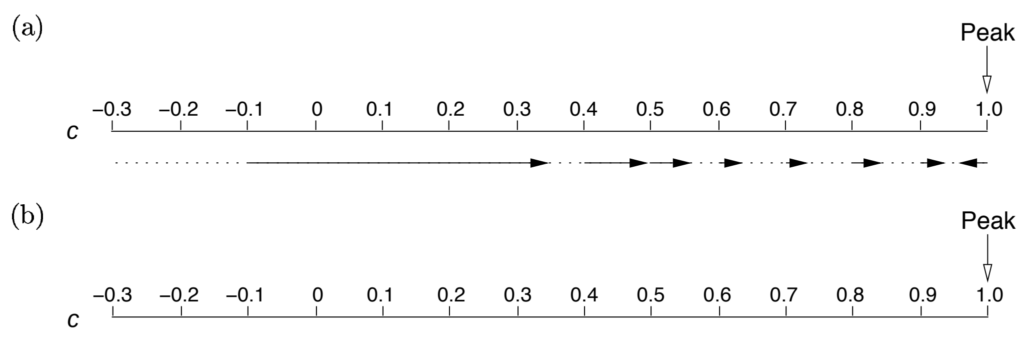

Figure 14: The dynamics of language change in high-density environments. In both the unstructured (a) and structured (b) environments there is overall movement towards $c = 1$, corresponding to the peaks of the distributions given in Figure 12. However, in the unstructured environment the speed of movement towards this point is lower, and there is a slight tendency to be repelled from $c = 1$. In the structured environment, movement towards the attractor is more rapid and there is no significant repulsion.

the signals associated with other meanings, and has to be observed to be learned. Consequently, compositional language in an unstructured environment is less stable through the transmission bottleneck.

6 Conclusions

We have presented a simulation model which demonstrates that compositional language can emerge from initially non-compositional language through purely cultural processes. Compositional language emerges when there is a bottleneck on cultural transmission — compositionality is an adaptation by language which allows it to slip through the transmission bottleneck. The advantage of compositionality is at a maximum when language learners perceive their world as structured — when the objects in the environment relate to one another in structured ways then a generalisable, compositional language is highly adaptive.

We are not, however, arguing that compositionality can be understood purely in terms of cultural evolution. The complex adaptive systems of learning and biological evolution still have a role to play. In the models described here, after exposure to a small set of utterances a learner's knowledge remains fixed. In the real case, however, an individual's knowledge and use of language is constantly changing and adapting, and this may impact on cultural evolution.

Similarly, we have offered no account of the biological evolution of the linguistic capacities of our simulated agents. Whereas the standard adaptationist model would hypothesise a complex, language-specific component of the brain designed to deal with compositionality, we make much weaker assumptions — a simple associative learning mechanism, in combination with a capacity to infer the communicative intentions of others, is sufficient to allow the cultural evolution of compositional language. It is not clear that either of these capacities are language specific, and an evolutionary account of their origins and development might be domain-general or exaptationist in flavour. The interaction between the evolution of this mental capacity and the ongoing cultural evolution of language is an exciting topic for future research, and computational modelling techniques will continue to be an invaluable tool in such endeavours.

References

[1] D. A. Baldwin. Infants' contribution to the achievement of joint reference. *Child Development*, 62(5):875–890, 1991.

[2] J. Batali. The negotiation and acquisition of recursive grammars as a result of competition among exemplars. In Briscoe [6], pages 111–172.

[3] P. Bloom. *How Children Learn the Meanings of Words.* MIT Press, Cambridge, MA, 2000.

[4] H. Brighton. Compositional syntax from cultural transmission. *Artifical Life*, 8(1):25–54, 2002.

[5] H. Brighton, S. Kirby, and K. Smith. Situated cognition and the role of multi-agent models in explaining language structure. In D. Kudenko, E. Alonso, and D. Kazakov, editors, *Adaptive Agents.* Springer, London, forthcoming.

[6] E. Briscoe, editor. *Linguistic Evolution through Language Acquisition: Formal and Computational Models.* Cambridge University Press, Cambridge, 2002.

[7] N. Chomsky. *Aspects of the Theory of Syntax.* MIT Press, Cambridge, MA, 1965.

[8] N. Chomsky. *Reflections on Language.* Pantheon, New York, NY, 1975.

[9] N. Chomsky. *Rules and Representations.* Basil Blackwell, London, 1980.

[10] N. Chomsky. *The Minimalist Program.* MIT Press, Cambridge, MA, 1995.

[11] M. Gell-Mann. Complexity and complex adaptive systems. In J. A. Hawkins and M. Gell-Mann, editors, *The Evolution of Human Languages*, pages 3–18. Addison-Wesley, Reading, MA, 1992.

[12] P. Glendinning. *Stability, instability and chaos: an introduction to the theory of nonlinear differential equations.* Cambridge, 1994.

[13] T. Hashimoto. Dynamics of internal and global structure through linguistic interactions. In J. S. Sichman, R. Conte, and N. Gilbert, editors, *Multi-Agent Systems and Agent-Based Simulation*, pages 124–139. Springer-Verlag, Berlin, 1998.

[14] J. A. Hawkins. A parsing theory of word order universals. *Linguistic Inquiry*, 21:223–261, 1990.

[15] J. R. Hurford. Nativist and functional explanations in language acquisition. In I. M. Roca, editor, *Logical Issues in Language Acquisition*, pages 85–136. Foris, Dordrecht, 1990.

[16] J. R. Hurford. Language learning from fragmentary input. In K. Dautenhahn and C. Nehaniv, editors, *Proceedings of the AISB'99 Symposium on Imitation in Animals and Artifacts*, pages 121–129. Society for the Study of Artificial Intelligence and the Simulation of Behaviour, 1999.

[17] S. Kirby. *Function, selection and innateness: the emergence of language universals.* Oxford University Press, Oxford, 1999.

[18] S. Kirby. Spontaneous evolution of linguistic structure: an iterated learning model of the emergence of regularity and irregularity. *IEEE Journal of Evolutionary Computation*, 5(2):102–110, 2001.

[19] S. Kirby. Learning, bottlenecks and the evolution of recursive syntax. In Briscoe [6], pages 173–203.

[20] M. Krifka. Compositionality. In R. A. Wilson and F. Keil, editors, *The MIT Encyclopaedia of the Cognitive Sciences.* MIT Press, Cambridge, MA, 2001.

[21] M. Piattelli-Palmarini. Evolution, selection and cognition: From "learning" to parameter setting in biology and in the study of language. *Cognition*, 31:1–44, 1989.

[22] S. Pinker. *The Language Instinct.* Penguin, London, 1994.

[23] S. Pinker and P. Bloom. Natural language and natural selection. *Behavioral and Brain Sciences*, 13:707–784, 1990.

[24] A. D. M. Smith. Establishing communication systems without explicit meaning transmission. In J. Kelemen and P. Sosík, editors, *Advances in Artificial Life: Proceedings of the 6th European Conference on Artificial Life*, pages 381–390. Springer-Verlag, Berlin, 2001.

[25] K. Smith. The cultural evolution of communication in a population of neural networks. *Connection Science*, 14(1):65–84, 2002.

[26] L. Steels. Constructing and sharing perceptual distinctions. In M. van Someren and G. Widmer, editors, *Proceedings of the European Conference on Machine Learning, ECML '97*, pages 4–13. Springer-Verlag, Berlin, 1997.

[27] L. Steels. The origins of syntax in visually grounded robotic agents. *Artificial Intelligence*, 103:133–156, 1998.

[28] L. Steels, F. Kaplan, A. McIntyre, and J. Van Looveren. Crucial factors in the origins of word-meaning. In A. Wray, editor, *The Transition to Language*, pages 252–271. Oxford University Press, Oxford, 2002.