Iterated Learning with Human Subjects: an Empirical Framework for the Emergence and Cultural Transmission of Language

Hannah Cornish, M.A. (Hons)

Linguistics and English Language
School of Philosophy, Psychology and Language Sciences
University of Edinburgh

August 2006

Submitted in partial fulfillment of the requirements of the degree of Master of Science

Declaration

I hereby declare that this thesis is of my own composition, and that it contains no material previously submitted for the award of any other degree. The work reported in this thesis has been executed by myself, except where due acknowledgement is made in the text.

Hannah Cornish

Abstract

The study of language evolution has benefitted enormously from the contribution made by computational simulations of the cultural transmission of language over the past ten years. However, we still have not explored or confirmed these findings empirically in a human population. This thesis presents a novel experimental method for investigating the emergence and cultural transmission of language under controlled laboratory settings. By integrating techniques from the modelling of iterated learning, with techniques used to investigate language acquisition via artificial language learning, a suitable empirical framework is created, opening up new avenues of research for understanding human language and culture.

Acknowledgements

I would like to thank the following for their support during this project:

- Firstly, to my supervisors Simon and Kenny – for all the common-room chats, ideas, hard-work, enthusiasm, blind-belief and motivating discussions.
- Secondly, to my friend Adam, who as always went beyond the call of duty to keep me sane and working hard when the whole of Edinburgh was in the Speigeltent living the high life.
- Thirdly, to all my fellow MSc-ers for being great fun all year – particularly Cyp, Seb, Steve, Karola, and Dave for helpful comments, crosswords, discussions and LaTeX fixes.
- Finally, I would also like to extent my gratitude to the ESRC who supported me for the duration of this Masters, my parents (who supported me before my Masters), and also to the University of Edinburgh Development Trust, who graciously awarded me a Small Projects Grant which funded these experiments and made it all possible.

Contents

List of Tables

List of Figures

CHAPTER 1

Introduction

The central question that any researcher on human language must ask is surely this: Why is language the way it is, and not some other way? Much is made of the fact that no other species on the planet has a communication system as complex and expressive as ours, and yet it is easy to take language for granted. As infants, we acquire it effortlessly. As competent speakers, we use it constantly. Even our private thoughts seem intrinsically shaped and enhanced by our unique system of symbolic communication, if only to the extent that some complex concepts and modes of reasoning can only be acquired through linguistic communion with others (Carruthers and Boucher, 1998).

Part of an adequate answer to this why-question is going to come from understanding how language came to be the way it is in the first place. In other words, from understanding how and why language evolved in our species and no other. However, the study of language evolution immediately presents us with a major problem, in that the phenomena that we wish to study is not actually visible to us (Christiansen and Kirby, 2003b). That is not to say that there is no evidence for linguistic evolution; we are surrounded by the evolutionary end-points of the process in the form of modern human languages. What we lack however, is an uninterrupted line of data that goes back to the incipient languages possibly spoken by our early hominid ancestors. Attempts to reconstruct language can only go so far (Fox, 1995), and as Hauser and Fitch (2003, p.158) point out, "linguistic behaviour leaves no fossils, and many characteristics of language appear unique to our species."

Properties of Language

A more fruitful way to proceed might be to take a closer look at the properties of modern language and discover if they can tell us more about how language evolution works. The first thing to note is that language exists at the intersection of three complex adaptive systems (as shown in Figure 1.1), each of which operates over a different timescale (Kirby and Hurford, 2002; Christiansen and Kirby, 2003a).

At one level, language is simply a system that is acquired by an individual over the course of that individuals' lifetime. At another level, we can see it as a developing organism in its own right, changing over the lifetime of the language. What the individual has learnt will influences the structure of the language itself, and hence what gets transmitted to be learnt by the next generation. Finally, the way that language is acquired will depend in part by the learning biases that have evolved biologically over the lifetime of the species as a whole. These innate learning biases will have in turn been fine-tuned as a consequence of the language structure, in order to better learn it. Thus each of these systems interacts with one another in non-trivial ways.

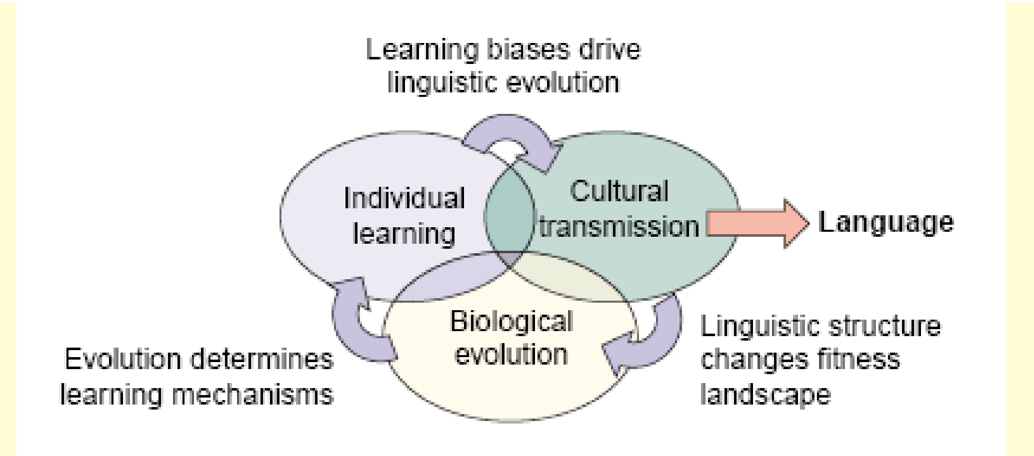


Figure 1.1: Language as a series of complex dynamic systems: this shows some of the ways in which these systems interact. (Taken from Christiansen and Kirby (2003a), with permission)
.

Cultural Transmission

We can describe the processes of how language is acquired by an individual (Bates et al., 2003), or how a biological trait can emerge via the process of Darwinian natural selection (Jablonka and Lamb, 2005) with a reasonable degree of accuracy – but considerably less is known about cultural transmission of language, or in fact, cultural evolution in general. In a recent paper, Mesoudi et al. (2006b) point out that if we compare the progress made in understanding biological evolution since Darwin's theory of natural selection was developed (Darwin, 1859) with the progress made in understanding cultural evolution, the study of culture is substantially lagging behind. This is unfortunate considering the mounting evidence that has been accumulated over recent years concerning the key role that culture plays in determining human cognition and behaviour (Tomasello, 1999; Richerson and Boyd, 2005).

As an example, let us consider what is actually meant by the term 'cultural evolution' According to Cavalli-Sforza and Feldman (1981, p.7), the term 'cultural' should be applied to "traits that are learned by any process of nongenetic transmission, whether by imprinting, conditioning, observation, imitation, or as a result of direct teaching." This clearly describes a wide and multi-faceted set of behaviours, all of which would benefit from a more thorough examination of the mechanisms giving rise to it.

Iterated Learning

In this paper I concentrate on just one type of non-genetic transmission process: observation. Language is not acquired by direct teaching[1], and nor is it learnt via imitation or conditioning as Behaviourists like Skinner (1957) once thought. Put simply, people learn their language by observing the linguistic actions of others. In this sense, language is a special kind of learning problem, where the output of one generation becomes the input for the next. This is better known in the literature on language evolution as iterated learning (Kirby and Hurford, 2002).

Most of what we know about iterated learning comes from computational models of human communication, where it has been shown to independently explain the emergence of universal features of syntax, such as compositionality (Kirby, 2000) or regularity (Kirby, 2001), without any simulation of biological evolution taking place. However, models are not perfect – their utility lies in the way that they can abstract away from the complexity of their subject matter (Cooper, 2002; Hurford, 2005). This has lead to some people arguing that real human populations would not act the same way as the model, that in general most simulations of language evolution contain "unrealistic initial conditions" which artificially limits the problem space, and makes the models work (Bickerton, 2003, p.86).

Up until now, there has been no empirical framework available in order to investigate whether this is true or not. However, if we look to the field of social psychology, there is an extant methodology that shares a

[1] I am of course referring only to a person's native language here – most people will be familiar with the idea of being explicitly taught a second language, although they may also be familiar with the all too common negative result that often accompanies it.

lot of similarity with the modelling of iterated learning: namely the *serial transmission chain* method devised by Bartlett (1932). In these experiments, information – usually contained as a written text – is given to a participant, who is then asked to recall it at a later time. This recalled version is then passed to the next participant and the process repeats itself, to form a chain of participants[2].

Traditionally this method has been used to investigate such things as how information (already encoded in natural language) changes according to some mental schema (Bartlett, 1932), social bias (Mesoudi et al., 2006a) or gender stereotypes (Bangerter, 2000; Kashima, 2000). The basic technique has also been modified in the past to look at such things as the convergence of a shared graphical representation scheme in humans (Fay et al., 2004), and even adapted to investigate the cultural conformity and spread of novel tool use in chimpanzees (Whiten et al., 2005).

I believe that with some further modifications these transmission chains can be adapted again to investigate how linguistic systems might arise in a human population; essentially implementing a human iterated learning model. This has the potential to offer a unique and novel empirical insight into the processes of language evolution, with the option of extending the paradigm to investigate other (non-linguistic) forms of cultural evolution that are transmitted via iterated learning.

Focus of the Current Paper

This thesis presents a novel experimental method for investigating the emergence and cultural transmission of language by using a series of artificial 'alien' languages that are learned and transmitted by human subjects over successive generations in the laboratory. The preliminary findings emerging from this new line of research will be analyzed in the context of the existing literature and results from computational simulations using the same parameters. In particular, it aims to show that the key finding to have emerged from the existing models – that languages evolve to become more learnable by their users – is upheld in a human population.

Road-map

It will begin in chapter 2 by exploring two different types of current literature that each contribute to the question of how language got to be structured the way it is: computational modelling of language evolution,

[2]This is just like in the game Chinese Whispers or Broken Telephone – see `http://en.wikipedia.org/wiki/Chinese_whispers` for how to play.

and empirical experimentation. In particular, it will present a thorough overview of the iterated learning model, before looking closely at recent work by Galantucci (2005) and Hudson-Kam and Newport (2005) – who present important new empirical work on the emergence of human communication systems and regularization in creole formation respectively. Additionally, work on social learning in primates will be touched upon, as it affords a unique insight into what an empirical framework for analyzing culture in humans might look like.

A new methodology which incorporates elements from all these areas will be outlined and justified in chapter 3, whilst chapter 4 will explore a series of experiments, the aim of which is to investigate some of the main predictions concerning when we can expect structure to emerge in language based on the computational simulations already discussed. Six studies will be detailed in total, before in chapter 5, a fuller discussion of the general implications of the data is explored. Finally, the thesis will be brought to a close with the inclusion of some proposals for extensions and future research.

CHAPTER 2

The Story So Far

2.1 The Iterated Learning Model

The main aim of developing the iterated learning model (ILM) was to explore the interaction that takes place between two distinct forms of linguistic representation (Kirby and Hurford, 2002): **E-Language** and **I-Language** (Chomsky, 1986). E-Language is basically the language that is externally represented to the world – the set of actual utterances. I-Language refers to the internalized representation of the language – stored in the brain as a pattern of neural connections. In order to get to I-Language from observing E-Language, people (or computational agents) must have some way of inducing a grammar based on what they have seen. Similarly, in getting to E-Language from I-Language, people and agents must have some method of production that does the converse (Hurford, 2002).

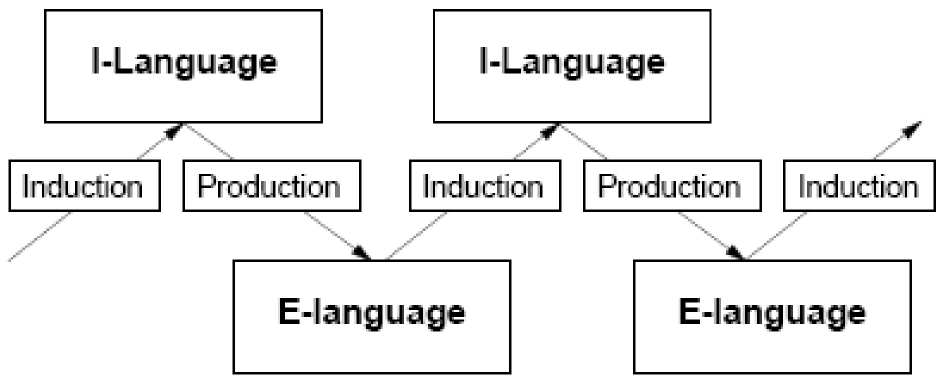


Figure 2.1: Transmission of I-Language to E-Language: as language is transformed between these two different forms of linguistic representation it is forced through a bottleneck. (Taken from Kirby (2001) with permission.)

2.1.1 Structure of the ILM

Just four components are required to implement an ILM:

- One or more learning agents;
- One or more adult teaching agents;
- A meaning space;
- A signal space

The task of the agents is to acquire a mapping between meanings and signals. The adult teaching agent has a random set of meanings that it must produce signals for. Usually the adults do not actually have a language system pre-specified at the start of the simulation, and so need to be provided with the ability to create an utterance string either by random invention, or by generalization from existing rules describing whole or partly 'known' meanings.

A typical simulation proceeds as follows:

1. An adult agent chooses a set of meanings to utter at random
2. Based on this set of meaning-signal pairs, the learner agent develops its own representation of the data.
3. After some specified amount of training (exposure to meaning-signal pairs) the learner becomes an adult, and the process repeats with a new learner.

There are a number of parameters in these models that can be varied: the size and structure of the meaning-space (Batali, 1998; Kirby, 2002a,b; Teal and Taylor, 2000; Zuidema, 2003); the production and induction mechanisms (Batali, 1998; Kirby, 2000; Tonkes and Wiles, 2002; Brighton, 2002); the population dynamics and structure (Smith and Hurford, 2003), etc. Perhaps the most fundamental parameter however is the bottleneck.

2.1.2 *The Role of the Bottleneck*

Every time E-Language is transformed into I-Language (i.e. every time language is acquired by anyone) a kind of filtering process takes place. As Deacon (1997, 110) explains: "The structure of a language is under intense selection because in its reproduction from generation to generation, it must pass through a narrow bottleneck: children's minds." Essentially what this is saying is that in order for a structure to survive the transmission process and become part of the I-Language, it must be learnable.

There are at least two ways in which a structure can increase its chances of survival and become more learnable. Firstly it could become more generalizable. One of the ways in which language can do this is by becoming *compositional* – structured in such a way that the meaning of the whole is composed of the meaning of its constituent parts and the way they are put together. In fact, this type of structure is one of the hallmarks of human syntax. Studies using the ILM explain the emergence of compositionality by virtue of the fact that a compositional element can appear in multiple different contexts, and thus maximize its chance of being learnt from the E-Language (Kirby, 2000).

The second way in which a structure can survive the transmission bottleneck is to just ensure that it is used frequently enough to make it likely to appear. This explains why it is that the ten most frequent verbs in English are all irregular (Francis and Kucera, 1982). They can afford to be, since they are used so often that it is guaranteed that a child will hear it and remember it in spite of it not fitting the regular pattern (Kirby, 2001). None of these results emerge in the models without the bottleneck being present. Furthermore, if there is no bottleneck (i.e. agents hear every single utterance in the language of the previous generation), compositionality will not emerge as the entire system can just be memorized. Conversely, if the bottleneck is too tight and only a few utterances get transmitted, the language will not be stable between generations.

2.2 Experimental Studies of Human Communication

So far, the picture we are building of language is one where the cultural transmission mechanism plays a central role in explaining the emergence of certain types of linguistic structure – creating selection pressures for language itself to evolve to become learnable by its human users. As such, we should be able to test this hypothesis experimentally.

2.2.1 The Artificial Language Learning Task

Christiansen (2000) suggested that one of the ways to do this is to use the Artificial Language Learning (ALL) paradigm. ALL has been used by researchers investigating language acquisition and the statistical learning abilities of humans (Saffran et al., 1996; Saffran, 2001) and non-humans (Fitch and Hauser, 2004)) for a number of years. Essentially it involves constructing a miniature artificial language with some desirable structural properties that we wish to investigate, training subjects on that language, and then testing them to see what they have acquired.

Building from computational modelling work that suggested word order universals in language come about because of language evolving to fit human sequential learning and processing constraints (Christiansen and Devlin, 1997), Christiansen (2000) devised two artificial languages that were either consistent or inconsistent with respect to the position of the head of a phrase in a sentence. The model predicted that the reason we do not naturally find languages with head-order inconsistency in the world is because they are too hard for us to learn. Results from the ALL experiment confirmed this, as subjects' performances were significantly worse on the inconsistent languages. An ALL task was also used to confirm similar modelling findings concerning the universal feature of subjacency in language (see Ellefson and Christiansen (2000) for more details).

ALL studies have also been used to investigate language formation in *creoles* – a type of contact language that emerges when speakers of multiple languages need to communicate (Thomason, 2001). The emergence of creole languages (Bickerton, 1981), along with recent studies of the emergence of a new sign-language found developing in the Nicaraguan deaf community (Kegl, 1994) are some of the rare occasions when we can observe a human communication system arise naturalistically. However, these naturalistic studies do not provide us with the experimental control to test our predictions.

Hudson-Kam and Newport (2005) used artificial languages to investigate the different roles that adults and children are hypothesized to play in creole formation. The phenomenon they chose to focus on was regularization – in particular, they posed the question of how languages with unpredictable variation (such as might be encountered in the early stages of creolization) become regular. They created two artificial languages that differed with regards to the presence or absence of a determiner within noun phrases. In one language the determiner was consistently present 100% of the time. In the other, it was only present 60% of the time. They taught both languages to a group of adults, and to a group of children between the ages of five and seven.

They used a series of tests designed to elicit both grammaticality judgements and production data in order to infer what the two groups had learnt based on their training, and found no significant difference between the performance of the adults and the children in the consistent group. In the inconsistent group however, they found a difference in that only the children regularized the inconsistent input. In other words they actively imposed a systematic pattern on the input to make it consistent. The way in which they did this varied from child to child – some left out determiners altogether, some included them at every opportunity, and one child only used a determiner with nouns in transitive sentences (Hudson-Kam and Newport, 2005). In contrast, the adult learners reproduced the inconsistency that they had heard.

Hudson-Kam and Newport infer from this that children play a crucial role in the formation of new languages: not necessarily 'inventing' it, but more likely acting to "regularize and stabilize the grammar of an emerging language" (p.185). This is perhaps not surprising considering the emphasis that the ILM places on the vertical transmission going from parent to child, as it is what the child is capable of learning, not the adult, that imposes the bottleneck on transmission[1].

2.2.2 Galantucci's Video Game Task

For a great many years, it has been assumed that studies such as those already outlined would be the closest we could ever come to empirically investigating the emergence of language. A paper by Galantucci (2005) has changed all that. In it, he describes an experimental study where pairs of subjects play a simple video game together. Each subject controls an agent that is situated within a virtual environment, initially composed of four rooms, each of which is uniquely identified by a shape in the center. The players' task is a simple one – to get their agents into the same room by making no more than one room change each. In order to do this, they have to somehow co-ordinate their movements. But there is a catch. Neither player can see the other player's agent until they are in the same room as one another, and the two players are prevented from communicating with one another using any language they know.

Instead they are given a digitizer pad which is hooked up to a computer upon which they may draw. On their computer screen they can see the effect that writing on the digitizer pad has, and also see what the other player draws. Again, here too there is a catch: the digitizer pad only records movement in the horizontal plane. In place of freely independent vertical motion, there is a constant downwardly scrolling motion, meaning that a left-to-right horizontal motion on the pad renders itself a top-left to bottom-right diagonal on the screen. This in effect prevents players from writing any alphanumeric characters, or even drawing any iconic images. However, in spite of these impediments, nine out of ten pairs managed to solve the game in less than 3 hours. Furthermore, they did this by negotiating a communication system together.

[1]Although the iterated learning model does not actually make any claims that the learner actually has to *be* a child – just inexperienced with respect to the stimuli being transmitted – as we will see in the later experiments all run on adult participants

By increasing the complexity of the environment (e.g. from 4 rooms to 9), and changing the task (find and capture the prey) Galantucci was able to demonstrate how these linguistic systems, once established, could adapt in response to the new demands being placed upon it (Steels, 2006). Whilst the resulting systems are difficult to analyse quantitatively due to their graphical form, a number of striking features appear. Firstly, the resulting systems are incredibly diverse, being based on either an arbitrary numerical system, the icon in the centre of the room, or on the layout of the game map. More interestingly, there was even indication that at least one pair (pair 7) had stumbled upon a method of re-using the symbols they created in the earlier games after the environment had expanded to 16 rooms, by coming up with a symbol meaning 'below' and a symbol meaning 'to the right of'. This combining together of different meanings into a new meaning for the whole is surely reminiscent of compositionality.

2.3 Experimental Studies of Social Learning in Primates

Clearly language is not the only skill that is culturally transmitted via observation. Much work in comparative biology has been focused on exploring the role of social learning in animal behaviour. The majority of this work has been focused on primates[2], trying to characterize the extent to which they resemble us culturally (Caldwell and Whiten, 2004; Whiten, 2005).

A lot of debate has gone into whether non-human primates can learn by imitating others, or by emulating others. Both involve observation and learning of new behaviours, but whereas imitation involves 'learning to do an act from seeing it done', emulation involves 'attending to the end-result in the environment' (Tomasello, 1998, p.704): not necessarily doing it the way it has been seen, but achieving the same environmental effect in the end. In order to try to distinguish between these two closely related types of observational learning, a series of experiments were devised that involved manipulating an 'artificial fruit' (Whiten et al., 1996; Caldwell and Whiten, 2004; Custance et al., 2001).

These experiments are of interest to us here as they explicitly attempt to observe *how* the skill (in this case, opening the fruit) has been transmitted from a demonstrator to an individual. The artificial fruit itself consists

[2] although see Rendell and Whitehead (2005) for a review of recent findings studying cetacean culture

of a box containing food that can be opened in one of two ways – by either pushing or twisting a bolt, opening a barrel latch by turning a pin and turning a handle, or opening a barrel latch by twisting a pin and pulling a handle. The different ways of opening the fruit can be seen as different learned cultural variants – even though all three methods could be discovered by a primate interacting with the box individually, if they only performed the variant that they were shown, we can be reasonably confident that this is because they have learnt this during the observation. All in all, the experiment was performed on eight different primate species (including human children).

Researchers have also managed to perform some similar larger scale studies in the field – for instance, introducing a new kind of nut to wild chimpanzees and observing the way in which the new nut-cracking technique that is required is eventually taken up by the group (Biro et al., 2003). These are perhaps even more interesting for the purposes of our discussion as they actively involve observing how the skill is transmitted *through the group*, rather than one-to-one.

Whiten et al. (2005) recently performed a more controlled experiment to investigate the cultural transmission of a novel tool within three different groups of captive chimpanzees. Each of the three groups was exposed to a new foraging task. In two of the groups, a high-ranking female was trained one-on-one with a human demonstrator on how to use the tool to get food from a 'pan-pipes' device in one of two possible ways: either by poking it with a stick, or lifting a catch with a stick. The third group acted as a control, and no expert was trained. The idea behind the experiment was to observe how different cultural variants of the task (poke and lift) would spread in the community when it was openly diffused like this, making it an interesting contrast with the direction of cultural transmission in the iterated learning model.

Instead of controlling it so that the expert teaches each chimpanzee in the group in turn (as would likely happen in the ILM), the entire group was allowed to observe their local expert obtaining food via the method over a seven day period, whilst the number of demonstrations each individual chimpanzee paid attention to was noted. During this time, the other group members could not access the pan-pipes. The task was then made available to all the chimps in the group over an additional ten-day period, and all attempts at performing the task were recorded. Firstly, 30 out of the 32 chimps in the experimental conditions managed to master the task – an impressive feat in itself. Secondly, although some chimps from both groups independently discovered the alternative method, the significant majority went on to conform to the 'normal' behaviour of their fellow group-members. The control group in contrast, despite intense interest in the task, failed to gain any food.

CHAPTER 3

The Missing Framework

3.1 Why the World Needs Alien Languages

We can summarize the preceding discussion as follows:

- Work on computational modelling suggests to us that language evolves to become learnable because of the way in which it is transmitted to users over generations (Kirby and Hurford, 2002);
- This theory gets some support from empirical studies investigating artificial languages (Christiansen, 2000; Christiansen and Ellefson, 2002);
- Further investigations of artificial languages show that children and adults may play different roles in regularizing languages (Hudson-Kam and Newport, 2005);
- Galantucci (2005) offers us an empirical way to investigate how communication systems emerge between pairs of speakers;
- Finally, we have looked at some studies that can show how a cultural trait such as foraging is transmitted through a population of non-human primates (Caldwell and Whiten, 2004; Whiten et al., 2005).

The question to pose now is, what does this all mean for the study of language evolution, and culture in general? The study by Galantucci (2005) shows us how languages can arise through negotiation in order to solve a specific task, but there is no attempt made here to investigate how that language may be transmitted beyond the pair (although this is certainly a topic for further research). Perhaps the most important contribution the work makes however is in showing that, against all expectations, it *is* possible to investigate language emergence in the laboratory.

The work being undertaken on diffusion chains in chimps is clearly more orientated toward exploring the transmission side, although it is not focused on language evolution. In addition, the cultural traits that have been explored using this method have all so far existed as simple binary variants – the box is opened by the bolt or the barrel latch; the food is obtained using the poke or the lift technique etc. – and as convincing a case as has been made for non-human primates having a rich cultural life and history (Whiten, 2005), there are limits to what a comparative approach can tell us about our own species[1].

Essentially we must conclude that there is still not a proper framework within which to examine how more complex cultural traits (such as language) emerge and are culturally transmitted. This is significant, not least because of the fact that there are many competing theories out there that would benefit from data gathered within an empirical framework designed to tap into the way in which complex nongenetic cultural traits are transmitted (Dawkins, 1976; Cavalli-Sforza and Feldman, 1981; Dennett, 1995; Aunger, 2000). The solution I propose is to integrate elements from all of the studies previously discussed in order to construct a new framework. In order to be useful and go beyond what we already have, this framework must have the following properties:

1. It must result in empirical data from human subjects being obtained
2. It must allow for experimental control and manipulation of variables
3. It must enable us to investigate a complex trait with many different variants open to cultural selection

In addition, it would be desirable for this framework to be flexible enough to allow us to explore the different modes of cultural transmission that have been identified, such as horizontal (peer to peer) and vertical (inter-generational) (Cavalli-Sforza and Feldman, 1981). To this end, a series of experiments were designed whereby artificial 'alien' language are constructed and given to learners, whose responses are given to the next learner, and so on. This simple idea, to essentially use techniques from ALL to implement a human ILM, has to my knowledge never been done before, despite having the potential to provide just the kind of framework we need to address the important questions facing cultural theorists and language evolution researchers.

The rest of this paper aims to test the viability of this suggestion by examining what, if anything, it can show us about language evolution. This will be done by attempting to confirm the main hypothesis to have

[1]Although it will still be very interesting to see what will be shown in the forthcoming work by Horner et al. investigating this kind of transmission chain in children.

emerged from the iterated learning of language, that language is an adaptive system in its own right that evolves to become learnable by its users.

3.2 Methodology

This section presents the methodology behind the studies discussed in the rest of this paper.

Generalized Overview

Each experiment took place using a computer, and consisted of 3 main steps repeated over ten generations of learners:

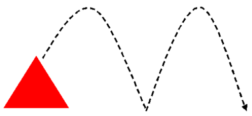
1. A subject is exposed to some sub-set of an artificial language consisting of a series of pictures (corresponding to meanings) paired with a string of letters (the signal). They are told that the string describes the picture in an 'alien' language, and it is their job to learn this language as best they can based on the pairings.
2. After this period of training, the subject is presented with a series of pictures without strings, some of which will be pictures they have encountered before during training, and some of which they will never have seen before. Their task is now to provide the correct description for each picture.
3. These pictures and responses then become the new set of training pairs for the next participant.

Participants were given both verbal and written instructions regarding the structure of the experiment (see Appendix A) but were not told the experimental aims, nor that their results would be given as input to another participant until after they completed the task. They were instructed to always type in a response, even if they were not sure of it, in order to "maintain good relations with the aliens". In addition, they were fully informed that they would undergo a series of three bouts of exposure to the language, followed by three periods of testing with optional timed rest periods. This was done both to break the task up for the participant and to let them practise typing in a response. Only their final responses were actually given to the next generation however.

It is important to note that there is no explicit training or instruction on the language beyond the presentation of the meaning-signal pairs, and no feedback was offered to the subjects by the experimenter during the task. Thus each person must discover the language anew based solely on the observed behaviour of the previous generation.

Structure of Meaning Space

The pictures contained a series of differently coloured simple geometric shapes, all engaged in some kind of movement activity. There were 27 of these 'meanings' in total, with three features along each dimension, as shown in table 3.1 below. The same meaning space was used in every experiment.



MOTION	COLOUR	SHAPE
spiral	red	square
bounce	blue	circle
horizontal	black	triangle

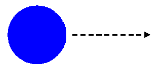


Table 3.1: Table showing the meaning-space structure used in all languages, with an example on either side.

Structure of Initial Alien Language

One of the key things Galantucci (2005) contributes in his experiment is the communication medium that ensure his subjects create an entirely novel language from scratch (Steels, 2006). Obviously, with a methodology as I have laid out here, an initial alien language of some sort is required for the very first person to learn. I propose to do this by giving the first learner a totally randomly constructed language. The initial languages were composed of nine syllables drawn from a set of forty, and then concatenated (with replacement) to form strings between 2 and 4 syllables long. These were then assigned arbitrarily to each of the 27 meanings. There are two reasons for doing this. Firstly, this is often how languages start off in the modelling research[2].

Secondly, I would argue that starting the languages off in this way is not so different from the way that creole languages often start from more simplistic *pidgin* languages. It is often said that a pidgin is no-ones native language; it clearly contains a blending or borrowing of extant structure from other languages, and yet it often becomes the point of departure for a fully fledged creole, whose eventual structure is almost unrecognizable based on its humble beginnings (Thomason, 2001). Thus, if anything, the

[2] This is a slight exaggeration – although the net result is a random language, more typically the models get the agent to produce this language themselves by using random invention. This is not a viable option in this case as it is hard to say what subjects would do if asked to produce a 'random language'. Maybe this is a topic for further investigation.

complete lack of structure initially should make it *harder* for subjects to find structure upon which to build a new language, not easier.

Participants

80 subjects (34 male, 46 female, mean age = 23;10, s.d. = 6;1) were recruited in total; 20 for two pilot studies, and a further 60 for the main study. Participants were deemed ineligible to take part if they: a) had studied linguistics beyond first year at University level[3] or b) suffered from dyslexia. There were no restrictions placed on the native-language of the subjects, although the majority of participants were native English speakers (73.75%), and all subjects reported themselves as being either competent or fluent in English. They were all paid the sum of £5 for their participation, and randomly allocated into one of the 8 language families based on when they were available to take part in the experiment.

Pilot Studies

Two pilot studies were performed with the objectives of a) refining the methodology to improve conditions for subjects, and b) ensure that sufficient data was collected to address the main hypothesis. Of most immediate concern was the fact that most ALL studies spend a great deal of time training subjects. For example, Hudson-Kam and Newport (2005) trained their subjects over six sessions of 10 to 20 minutes each, spread over 9 days. Due to the fact that we were getting our training data from the output of previous participants however, a similar regime was impossible to implement here – in this time-scale, and provided there were no drop-outs by participants, each language family would take three months to complete. With that in mind, it was decided that each experiment should last no longer than an hour. It was therefore important to know how much of the language could be acquired in this time.

Following the first pilot study, the decision was made to collect signals for all 27 meanings on the last test, instead of just those that were to be given to the next generation. This was done because in calculating intergenerational accuracy scores, we were often comparing a word created in generation 10 with a word last uttered in generation 4, which seemed no longer representative of the language that generation 10 had built up

[3]These restrictions were put into place as a precaution only, as it was felt that these people would be tempted to analyse the task in a different way to the general public.

since. This did increase the amount of time spent testing, although post-test questionnaires performed on subjects exposed to this longer test condition did not highlight it as a concern. Interestingly enough the vast majority of the participants (17 out of 20) reported that they were unaware that they were being tested on items that they had not seen. Most importantly overall however, subjects levels of recall for the seen items suggested that they were all capable of learning the language in the time they were exposed to it, and although the task was uniformly felt to be 'difficult', subjects agreed that it got easier by the third round of training, and even reported that the task was enjoyable.

Procedure

Two basic conditions were initially identified, based on the notion of the size of the bottleneck:

- In the **50% condition** subjects were trained on 14 meaning-signals drawn randomly from the language. These 14 items were called the *seen* set, and the 13 remaining items were called the *unseen* set. During the test phases, they were tested on seven meanings from the seen set and a further seven from the unseen set - both of which were sampled randomly. In the final testing run, the extra 13 items (7 seen, 6 unseen) were tested on directly after the usual test-set. This was done to ensure that any influences of fatigue would not directly affect the language composition to be presented to the next generation.
- In the **75% condition** the same basic procedure was followed, only this time the subject was trained on 20 meaning-signals which formed the seen set. During testing, they were given the entire contents of the unseen set (7 items) plus 13 items sampled randomly from the seen set, with a final test set which includes the remaining 7 unseen items.

Exposure to the stimuli was timed using a computer: each meaning-signal pair appeared for exactly six seconds, with the signal preceding the meaning by one second[4]. Each meaning-signal pair in the training (seen) set was randomized and given to the participant twice before a test and a break. This break was optional but was timed to ensure it was no longer

[4]This was done in part to share parsimony with the order of events in the ILM, but also as a result of a post-test questionnaire following the first pilot.

than two minutes. The exact schedule for both sets of experimental conditions is as follows:

- Training on seen set (x2)
- Test
- Break
- Training on seen set (x2)
- Test
- Break
- Training on seen set (x2)
- Extended Test (of entire language)
- End

Chapter 4

The Human Iterated Learning Model

This chapter describes the results of running the human iterated learning model. In order to uphold the central hypothesis that language evolves to become learnable, it needs to be shown that not only does subjects' performance on the seen items increase, but also their performance on the unseen items. In order to do this we need to operationalize some measure of performance. Additionally, it might be useful to have some way of describing the structure of the language itself in quantitative terms.

Subject Performance on Language

The performance of each subject was judged by comparing their language with the language spoken by the previous generation – in other words, by measuring how faithfully the subject was able to reproduce their input. As such, it can also be seen as a measure of language stability.

In order to calculate this inter-generational performance, we need to start with some way to measure how similar two output strings are. As we do not have any control of the strings that a subject will output, this similarity measure must be tolerant to variations in the lengths of both strings, and so Levenshtein Distance (Levenshtein, 1966) was selected. The way it works is by calculating the minimum number of insertions, deletions or substitutions required to turn one string into another. For example, if we were to try to turn the string 'nanimi' into 'hanim' we would have to make one substitution ('n' to 'h') and one deletion ('i' to '$\emptyset$'), resulting in a distance of 2. In order to make useful comparisons with this figure however, we need to normalize it using Equation (4.1), where s_1 and s_2 are the strings we are comparing.

$$nLD(s_1, s_2) = \frac{LD(s_1, s_2)}{\max(len(s_1), len(s_2))} \tag{4.1}$$

Accordingly, identical strings score a value of 0, and maximally distinct strings score a value of 1. These figures can be summed for every utterance in the language, and then averaged to provide a single figure showing the stability of the language/performance of the subject over the whole language, or we can calculate performance on seen and unseen items independently.

Language Structure

When we describe something as structured, we are saying that there is something consistent and special about the way it is arranged. When we talk of language being structured, we usually mean that there is some regular relationship between a form and its corresponding meaning. In a compositional system, this relationship can be described as structure- or neighbourhood-preserving, reflecting the fact that neighbouring meanings in the meaning space will map to neighbouring signals in the signal space (see Kirby (2001) and Smith (2003) for more details).

One way to measure the amount of correlation between the relative distances in the meaning and signal spaces (and hence discover if similar signals map to similar meanings) is to examine the pairwise differences between them. The equations in (4.2) and (4.3) define an estimate of language structure over an entire language of n meaning(m)-signal(s) pairs. To calculate $E(O)$, the average Hamming and Levenshtein distances ($\overline{HD}$ and $\overline{LD}$ shown in (4.3)) are required[1]. The average Levenshtein distance uses every pair of meanings, excluding comparison against itself (i.e. n^2-n pairs), and an identical number of signal pairs for the Hamming distance, hence the denominator of $n(n-1)$. $E(O)$ is then simply Pearson's correlation coefficient with the summations used to calculate the mean differences for both the Hamming and Levenshtein distances.

$$E(O) = \frac{\sum_{i=1}^{n}\sum_{\substack{j=1\\j\neq i}}^{n}\left(HD(m_i,m_j)-\overline{HD}\right)\left(LD(s_i,s_j)-\overline{LD}\right)}{\sqrt{\left(\sum_{i=1}^{n}\sum_{\substack{j=1\\j\neq i}}^{n}\left(HD(m_i,m_j)-\overline{HD}\right)^2\sum_{i=1}^{n}\sum_{\substack{j=1\\j\neq i}}^{n}\left(LD(s_i,s_j)-\overline{LD}\right)^2\right)}} \tag{4.2}$$

[1]The Hamming Distance is another string edit distance measure which works by calculating the lowest number of substitutions required to transform one string into another. As such, it requires the two string lengths to be equal, which for the meaning space instantiated here, is the case.

$$\overline{HD} = \frac{\sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} HD(m_i, m_j)}{n(n-1)}$$

$$\overline{LD} = \frac{\sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} LD(s_i, s_j)}{n(n-1)} \tag{4.3}$$

This should result in a value between 1 and 0[2]. (Smith, 2003) argues that a value of approximately 1 represents a fully compositional system, whereas a value of approximately 0 represents a holistic system (one where the meaning of the whole is not a function of the meaning of its parts). In order to confirm that the random initial languages that were created by the method described in section 3.2 are in fact holistic, 1,000 were generated and their language structure estimated. This resulted in a mean of 0.0006, and standard deviation of 0.04 – well within the expected range. This shows us that the chances of us stumbling upon a compositional language via the random generation process is vanishingly small.

4.1 Experiment 1: Random Initial Language

In this experiment two language families were created with different initial languages drawn at random. Family 1 were trained with a 50% bottleneck, whereas Family 2 had a 75% bottleneck.

4.1.1 Learnability Increase

To test whether the languages changed to become more learnable or not, subjects' performance on seen and unseen items was calculated (see Figure 4.1). In both bottleneck conditions the nLD of the seen items falls dramatically over generations, eventually settling somewhere around zero. Examining the graphs of the unseen items, we can see that these also fall to zero, although not as quickly nor as smoothly as for the seen items. In addition, the language with the bigger bottleneck seems to stabilize faster than the one with the smaller one. This shows several things: firstly, participants who encounter the language at a later generation are finding it easier to

[2]Actually, the Pearson's Correlation produces a number between 1 and -1, but it is quite hard to visualise what a language with a negative correlation would look like. It should be one where similar meanings are associated with dissimilar signals, and vice versa.

learn than those at earlier generations, and secondly, the bottleneck may be affecting the speed at which learners converge on a stable language.

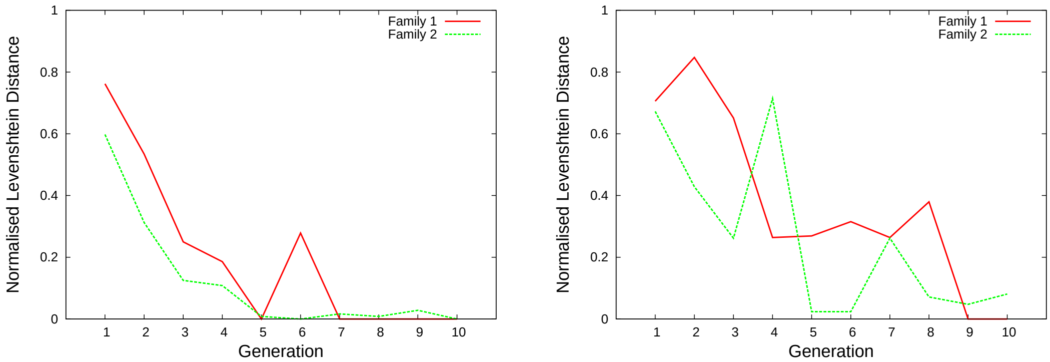


Figure 4.1: Graphs showing performance on seen (*left*) and unseen (*right*) items in family 1 and 2: the normalized levenshtein distance of both sets of items falls over time, indicating the language is becoming easier to learn.

It is worth considering in more detail why it is that the subjects start to perform better on the unseen items with time. If the language remains random, we would expect this score to be uniformly poor across generations. The only way for a subject to 'know' what the correct description is for a meaning they have not seen is to somehow infer it based on the relationships between the meanings and signals that they *have* seen. If we look more closely at the languages themselves (see Appendix B.3 and B.4) it appears that this increase in language performance is happening at the same time as a vast reduction in the number of distinct strings in the languages: although both language families start out with 27 distinct strings, by generation 10 this figure has fallen to 5 in each case. This means that the meanings are now underspecified with respect to the signals in the language – instead of a one-to-one mapping, we have ambiguity.

4.1.2 *Structured Ambiguity*

Looking at the languages at the final generation however, it appears that this ambiguity is not random. We find that we can describe the languages in terms of rules. For example, in Family 1 everything that moves in a spiral is called 'poi', and everything that moves in a horizontal line is called 'tuge'. In addition to this, there is a three-way distinction for bouncing shapes: 'tupim' refers to bouncing squares, 'tupin' to bouncing triangles, and 'miniku' to bouncing circles (see table 4.1). A similar system appears to be at work in family 2 (table 4.2), where the signals also seem to describe a maximum of two of the three dimensions of meaning. As one participant (Family 1, Generation 9) put it: "the aliens don't seem to care about colour". For want of a better term, we will call this phenomenon 'Structured Ambiguity'.

	SQUARE	CIRCLE	TRIANGLE
SPIRAL	poi	poi	poi
HORIZONTAL	tuge	tuge	tuge
BOUNCE	tupim	miniku	tupin

Table 4.1: Structured Ambiguity in Family 1: This table shows how the 5 words in the language are divided up in terms of motion and shape. Colour is not reflected at all in this language.

	BLACK	RED	BLUE
SQUARE	horare	honare	honare
CIRCLE	gomia	gomina	[**gomina**;horare]
TRIANGLE	kakawa	kakawa	kakawa

Table 4.2: Structured Ambiguity in Family 2: This table shows how the 5 words in the language are divided up in terms of shape and colour. Motion is not reflected at all in this language. Items in bold font indicate they appeared with greater frequency.

As has already been mentioned, the number of signal types has fallen dramatically in these languages. As this happens, there is conceivably a possibility that the arrangement of signals to meanings could have occurred by chance, based on simply replicating the various frequencies of the seen items probabilistically. In order to prove otherwise, a montecarlo test was performed[3]. The idea behind this test is to ask what a totally naive agent, who had no access to the meanings, would do given the same data that the human was trained on.

The agent is given the training data for the language at every generation, and then asked to produce the same number of utterances as the human. As it cannot make a decision based on the meanings, it selects a word probabilistically from the training set based on their relative frequencies. For example, imagine a training set consisting of just two different words, A and B. Now imagine that A appeared in the training set 8 times, and B only appeared twice. The naive agent would pick A 80% of the time, but there is still a chance (20%) that it will pick B. Once it has picked the correct number of utterances, the nLD was calculated for the agent's choice.

[3]The code for performing this test was graciously provided by Kenny Smith.

This procedure was repeated 1000 times, and the 95% confidence interval plotted as an error-bar, before being compared against the actual figures obtained for the human, as illustrated in Figure 4.2.

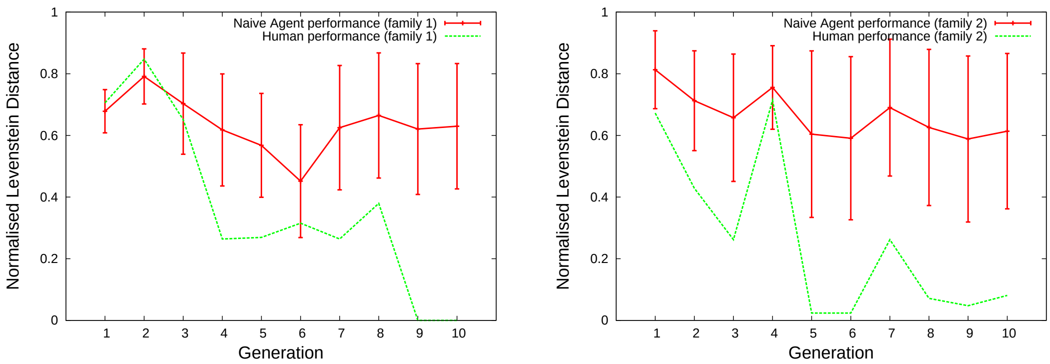


Figure 4.2: Performance of Naive Agent on Family 1 and 2: from this graph we can see that the human score is significantly different from that of the naive agent, trained on the same data. This indicates that the humans are inferring structure in the language.

We can see from this graph that whilst the human performance increases over generations, the naive agent's performance stay relatively stable. Every point where the human performance level is below one of the naive agent's error-bars represents a statistically significant difference in the relative scores obtained ($p<0.05$). What this shows is that the humans are definitely *not* following a similar strategy to the naive agent and just uttering words based on their probabilistic frequencies (at least not at later generations). They do better than the naive agent because they are guided by some underlying structure between signals and meanings, one that allows them to correctly infer a signal given a meaning, even in the absence of training on that specific meaning.

4.1.3 Discussion of Experiment 1

The most important thing of note to emerge from the first experiment is that we have confirmed the main learnability hypothesis from the computational simulations. The second thing of note was the way the languages changed to bring this about. Instead of becoming structured via compositional means, the languages became structured by underspecifying parts of the meaning space. This is actually quite an exciting finding, especially when we consider that it is a fairly well-established fact, both within and without the modelling community, that language learners appear to have a preferential bias towards having a one-to-one mapping between surface forms and underlying meanings (Slobin, 1977; Pinker, 1984; Markman and Wachtel, 1988; Smith, 2003). Why then, would learners be over-riding this natural instinct?

The interesting thing is that we find similar behaviour arising within the ILM's themselves. All of the many different computational instantiations of the ILM make explicit attempts to encourage one-to-one mappings to emerge (see Smith (2003) for a summary of the biases and methods used). In the literature this is usually discussed in terms of homonymy and synonymy[4]. As De Beule et al. (2006, p.467) explain:

> "The challenge in constructing artificial communication systems is to avoid homonymy (having words that have more than one meaning) and synonymy (having several words for the same meaning). If a communication contains a too high degree of homonymy or synonymy it can not be used effectively."

When the computational simulations are run without any constraints on producing distinct strings, they degenerate in a very similar way to what we see here – the number of strings, even if initially high in the original language, plummets and in most cases results in the model just producing a single utterance (Cornish, 2005). The fact that the human subjects do not let their language degenerate that far may be testament to the fact that they are still under the influence of their endogenous biases.

One way to look at the problem is to see the effect that the ambiguity is having on the actual task. We know that it is making the language easier to learn, so presumably the underspecification is somehow adaptive. Both family 1 and 2 become underspecified in the same way – by losing the ability to describe one dimension of their meaning space. This may not sound like much stated in these terms, but in fact, what that does is reduce the meaning space from 27 items to just 9. It is possible then, that the subjects are actually finding the task too difficult, and this is one way of alleviating the pressure of co-ordinating a complex meaning space. Of course, this is

[4]There is a reason why I have not just labelled the ambiguity found in this study as either homonymy or synonymy – the two are notoriously difficult to distinguish (Lyons, 1995; Pinkal, 1995), and this instance is no exception. On the one hand a case could be made for homonymy as we clearly have words such as 'poi' in Family 1 appearing for many different meanings. If what is being suggested here is correct however, this structured ambiguity is actually involving a reclassification of the *meaning-space* of the language - hence it is perfectly correct that all these instances of 'poi' should be named thusly, as they all refer to just one meaning: RED. If we gave the subject nine identical pictures, we would expect to get nine identical answers. In this instance, we cannot be sure that the subject is not interpreting the pictures as equivalent.

not to say that this occurred as a result of a conscious decision by any single participant. Once two identical words appear in the input to the next language learner the trend is only going to persist and snowball.

In the light of this finding then, would it even be possible to maintain a compositional system if one managed to emerge? And secondly, could a compositional system emerge if similar steps were used as in the models to curb ambiguity? These questions will both be addressed in the next two experiments.

4.2 Experiment 2: Compositional Initial Language

In this experiment, another two language families were created with different initial languages that were designed to be compositional from the outset. These languages were created by drawing nine syllables from our set of 40, and arbitrarily assigning them to individual components of meaning, as table 4.3 demonstrates for Family 4. In order to create a sentence, these words had to be combined together according to a strict word order – motion, followed by shape, followed by colour. This scheme was deliberately chosen so as the verb (motion) appeared first in order to make it operate differently compared to English.

SPIRAL	lu	TRIANGLE	ki	RED	na
HORIZONTAL	ka	CIRCLE	no	BLUE	ni
BOUNCE	po	SQUARE	me	BLACK	we

Table 4.3: Initial Language for Family 4

Family 3 were trained with a 50% bottleneck, while Family 4 had a 75% bottleneck.

4.2.1 *Bottleneck Affects Maintenance of Compositionality*

By plotting our measure of language structure over generations, and also examining subjects' performance on unseen items (Figure 4.3) we can see that the two languages, although start off in an identical position, eventually end up being rather different. The compositional system of Family 3 starts off being fairly close to that of Family 4, but by generation 8 the system is very close to being entirely random. In contrast, Family 4 (which has the larger bottleneck) seems to retain a high degree of language structure throughout, dipping a little at generation 7, but eventually ending up at a reasonably high figure of 0.8.

This difference in behaviour continues with regards to the performance on unseen items. In a compositional system, we would expect this value to be quite low, reflecting the fact that it is not necessary to actually see every item, just enough examples to be able to generalize the rules. This is what we initially find in both languages, but whilst Family 3's score increases (in marked contrast to the eventual end-point of the families in Experiment 1 – Figure 4.2), Family 4's score remains comparatively lower, although there is certainly a peak at generation 7 before the language eventually settles.

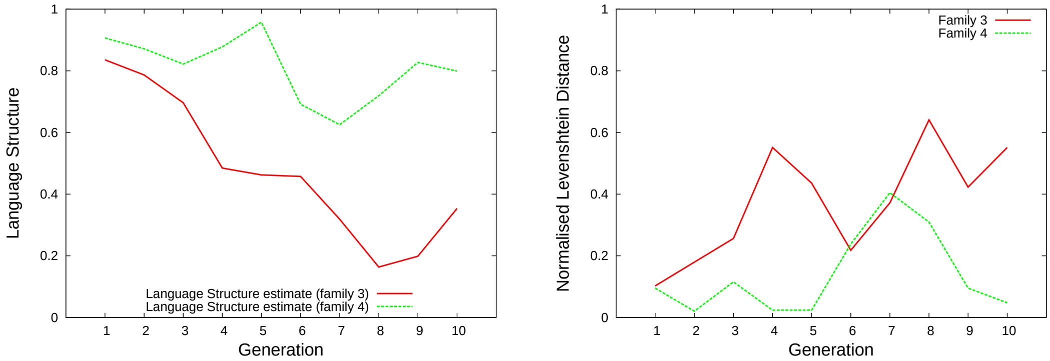


Figure 4.3: Graphs showing language structure (left) and performance on unseen items (right) in family 3 and 4: here we can see that whilst family 4 maintains a high degree of language structure throughout generations, family 3 does not. Additionally, the performance of subjects in family 4 in terms of unseen items is also better than that of family 3.

If we examine the languages in their final state (Appendix B.5 and B.6) we can see that although the language structure score is not quite at 1 for Family 4, the language does indeed show signs of compositional structure (table 4.5), albeit with some errors. These errors were of a specific type, in that the 'syntax' of the system was never compromised. All of the strings produced would have been grammatical according to the language, but just referring to a slightly different meaning. For example, 'kasawe' was produced twice; once correctly to refer to horizontal-circle-black, but also once to say spiral-circle-black.

In contrast to this, at generation 10, family 3 appears more chaotic (table 4.4). There are some local rules that are apparent – such as a marked similarity between BLUE and BLACK which contrasts with RED, or between red-squares and red-circles in contrast to red-triangles – but the patterns are clearly not representative of the structured ambiguity previously described, nor are they compositional. This perhaps explains the fact that they are not capable of being transmitted faithfully between generations, as even if the seen items are memorable, the unseen items must be invented anew by every generation.

	BLACK	BLUE	RED	
	megopu	megome	haheki	HORIZONTAL
CIRCLE	megome	megopu	hahepu	BOUNCE
	magopu	megopu	haheto	SPIRAL
	magoho	magohe	hahepu	HORIZONTAL
SQUARE	magome	magohe	haheto	BOUNCE
	magomo	magopu	hahepu	SPIRAL
	megomo	megoho	wikipo	HORIZONTAL
TRAINGLE	megoho	megohe	wikipu	BOUNCE
	megoho	megohe	wikipu	SPIRAL

Table 4.4: Analysis of language at generation 10 of Family 3. This table shows that whilst there are a few regularities here, the language is largely chaotic, with no systematic pattern accounting for the appearance of ambiguous signals.

4.2.2 Compositional System Tolerates Change

As already mentioned, Family 4 maintains its compositional structure right up until the final generation, although Figure 4.3 does show that there is some amount of instability in this, centred around generation 7. If we examine the language closer, we can see that this is because the 'words' comprising it have actually changed. It is still compositional, and is still reminiscent of its initial form, but five of the nine syllables have been altered (table 4.5).

SPIRAL	**la**	TRIANGLE	**ka**	RED	na
HORIZONTAL	ka	CIRCLE	**sa**	BLUE	**ne**
BOUNCE	po	SQUARE	**ma**	BLACK	we

Table 4.5: Language at Generation 10 of Family 4. Items shown in bold were items that had changed from the initial system learnt by the first generation. Note that due to its compositionality, the language can tolerate the syllable 'ka' having a dual significance.

This happened gradually, but during the time characterized by fluctuations shown in Figure 4.3. For instance, the word 'no' meaning CIRCLE remained stable right up until generation 7, at which point there were four different variants. The change over time for this one component is shown in table 4.6. It is interesting to note that the introduction of the 'sa','so, and 'se' variants occur in spite of the subjects at no point seeing any syllable starting with an 's'. This kind of random innovation is typical, and seen in every family.

Gen 6	Gen 7	Gen 8	Gen 9	Gen 10
no	**no**	**sa**	**sa**	**sa**
	sa	so	so	
	so	na		
	se			

Table 4.6: Competition amongst signal variants for CIRCLE in Family 3: Item is bold were the most frequent. A possible reason for the shift from 'no' to 'sa' is that syllables beginning with 'n' stat to feature more prominently at the *ends* of words, not medially.

We can speculate as to why 'no' lost its place in the language by examining the other syllables that begin with 'n' that feature in generation 7's response. In addition to 'no', we also find the syllable 'ni', 'ne' and 'na' all occurring at the end of the string, representing colour. In fact, from Appendix B.6 we can calculate that the person at generation 8 saw 10 instances of 'nX' at the end of a word, and only 4 instances of 'nX' medially. Against this number of 4, they saw 3 instances of 'sX' medially in the context of CIRCLE, but also an additional 2 times used 'incorrectly' against other shapes. This could easily have led the person at generation 8 to form some vague rules along the lines of *sX is related to shapes* and *nX is related to colour.*

4.2.3 Discussion of Experiment 2

Here the main finding is that compositionality can be maintained by the subjects, provided that the bottleneck is large enough. One major difference between the simulated ILM's and the experiment is here there are really two bottlenecks being imposed. The first is the one that we have control over – what proportion of the language is seen by a subject. However, the additional bottleneck that I am speaking about concerns the subjects own memory. Although nine syllables seems like a small number, it was clear that the subjects taking part in this experiment found the task a lot harder than subjects in the first experiment. This can be seen by examining the average response-time for each test item over the 4 families studied so far (Figure 4.4).

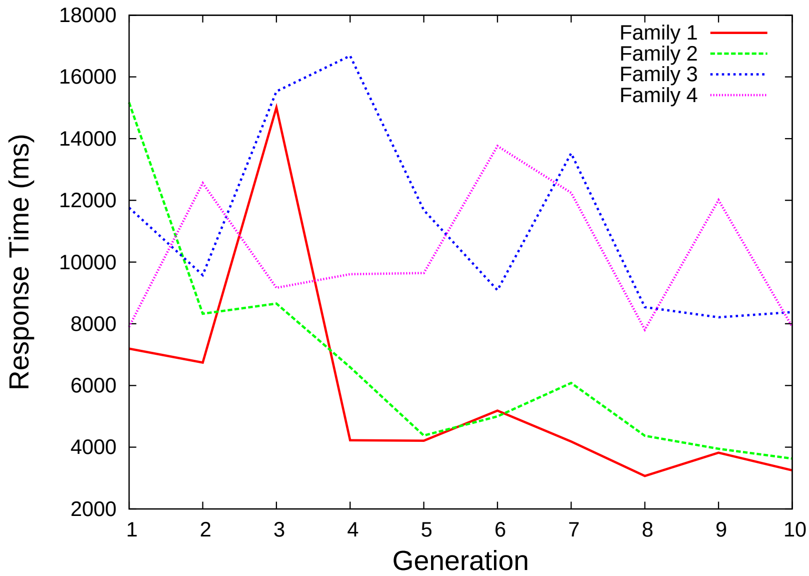


Figure 4.4: Graph showing average response time (ms) over all test items in families 1-4. This shows a marked contrast between the response times of families 1 and 2 at later generations, and the response times of families 3 and 4 at later generations. This indicates that learners of the initially compositional languages found learning those languages difficult throughout.

This graph shows that whilst the time it took to respond to the systems with structured ambiguity fell, the system that maintains its degree of compositionality, and the system that loses its compositionality pattern differently. It is perhaps conceivable that the reduction in response time is in some way a reflection of the fact that the number of distinct items decreases in families 1 and 2. If this is the case however, then the relationship is clearly not a straightforward one: if we look at the number of distinct items in family 3, we find that at generation 6 it falls to just 4 – lower than both family 1 and 2 – and yet the response time is much greater[5]. Additionally, although it was often clear from a person's responses (and the verbal descriptions they provided at the end of the test) that they understood the language, not a single subject in Family 4 managed to get 100% accordance with the previous generation. Clearly formulating words in the compositional language required a great deal of concentration. In this case, having a larger bottleneck is helpful because the extra examples and training it affords presumably make it easier to remember the elements and the way they are put together.

[5]Of course, each data point here represents a single subject, and so differences should be treated with caution.

Both this experiment and the preceding one seem to suggest that structured ambiguity is arising because it is an easier solution (or in other words, a better adaptation) to the pressures that the learning task is forcing upon the subjects. It will be interesting then to see what occurs in the next experiment when the ability to make the languages underspecified is taken away from the subjects.

4.3 Experiment 3: Ambiguity Filtering

As already mentioned in section 4.1.3, computational simulations of iterated learning have to impose certain biases in the agents in order to preserve one-to-one mappings between meanings and signals. These have been implemented in various ways – by penalizing or rewarding an agent with costs for avoiding ambiguity (Batali, 2002), by using an architecture that simply does not allow it (Hare and Elman, 1994; Batali, 1998; Kirby and Hurford, 2002) or even only allowing the agent to learn from and remember the first example it sees of any given meaning-signal pair (Kirby, 2002a). This experiment involves attempting to enforce one-to-one mappings with the latter approach.

Obviously it is physically impossible to prevent a human subject from learning from all and any examples it is given, but the methodology does allow us to filter out any ambiguity before it even reaches the next generation. The way in which this was done was to examine the output from the previous generation in order to identify instances where the same string appeared for different pictures. If found, only the very first instance of the string and the meaning associated it would be presented to the next generation, and the other meaning-signal pairs discarded; these items were not replaced with alternatives.

Two families were created with random initial languages as before. Family 5 were trained with a 50% bottleneck, and Family 6 with a 75% bottleneck, although with the filtering in place, the actual size of the bottlenecks fluctuated between individual generations depending on the number of ambiguities that were introduced by the previous subject. In any case, the number of items requested in the test phase remained equivalent to the bottleneck size. Whilst this did mean it was entirely possible to be trained on fewer items than you were asked to reproduce, it avoided getting to a situation where the subject was trained on just a single item.

4.3.1 No Increase in Learnability

When we look at the performance on both seen and unseen items (Figure 4.5) we find the subjects in family 5 seem to do better than family 6. Although nLD seems go down for both over time in the seen condition it still remains much higher than the random initial languages in experiment 1. Similarly, when we look at the performance on the unseen items, we see that again, family 5 and 6 do poorly in comparison to languages in experiment 1. Here however there seems to be a noticeable difference between the families in the different bottleneck conditions: family 6 barely shows any performance increase over time at all.

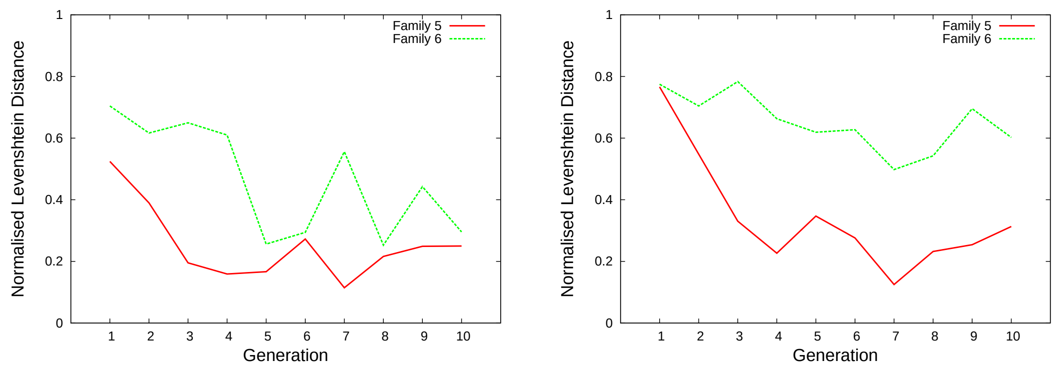


Figure 4.5: Performance on seen (*left*) and unseen (*right*) items for Family 5 and 6. Both families fail to show the same kind of performance improvements over time as witnessed in earlier experiments. Family 6 in particular does poorly in terms of unseen items, suggesting that this language is not becoming easier to learn over generations.

This indicates that filtering out the ambiguity is preventing the languages from stabilizing. In a way, this is what we want in order to confirm that the structured ambiguity in Experiment 1 is a positive adaptation, but it is curious as to why there should be such a difference between the two different bottlenecks. Family 5's performance is at the very least, questionable, whereas Family 6 does not seem to be becoming easier to learn over generations at all. We can examine what is going on in more detail by looking at how our naive agent would perform on the same signal data given to both families (Family 4.6).

This shows quite strikingly that, certainly for family 6, the subjects and the naive agent appear to be doing the exact same thing – making responses based probabilistically on the relative frequencies of the words they have already seen. This is highly indicative of the fact that there is no structured relationship between meanings and signals in this language

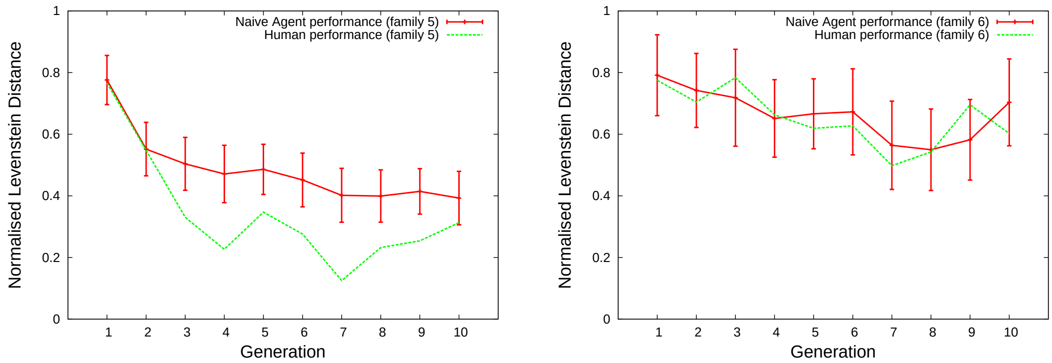


Figure 4.6: Performance of Naive Agent on Family 5 and 6: here family 6 is clearly doing no better than the naive agent on the same data. From this we can infer that the human subjects are following the same strategy. Interestingly enough, the score for the naive agent actually improves over time for family 5, meaning that the strings in the language are evolving to become more similar with time.

at all. Looking to family 5 however, it is interesting to note that the naive agent itself seems to be doing slightly better at the language over time, although still not quite to the same extent as the humans. What this seems to suggest is that the language is changing in such a way as to make it more likely to score well by just picking a seen item at random. This will be discussed in more detail later.

4.3.2 Emergent Compositionality

In the models, this type of filtering process has been shown to facilitate the emergence of compositionality. Given the negative results pertaining to the learnability of the languages here, it seems unlikely that the systems have become compositional. In fact, examining the final states of both languages, they clearly are not (Appendix B.7 and B.8). It is quite difficult to describe them in terms of any of the systems we have already discussed; they are most reminiscent of the chaotic system seen in Family 4. Nevertheless, running the estimate of language structure reveals that whilst neither language attains full compositionality, it is not true to say that there is no structure in the languages. In family 5 in particular, we see two peaks – at generations 4 and 7 – that look like they may have indicated structure was in the language at some point.

If we examine these in closer detail we discover a system that, whilst not being fully compositional, at least bears the hallmarks of it. To explain, the language at generation 4 is shown in table 4.7. Examining it, we can clearly see that each string is broken into at least two parts. The endings

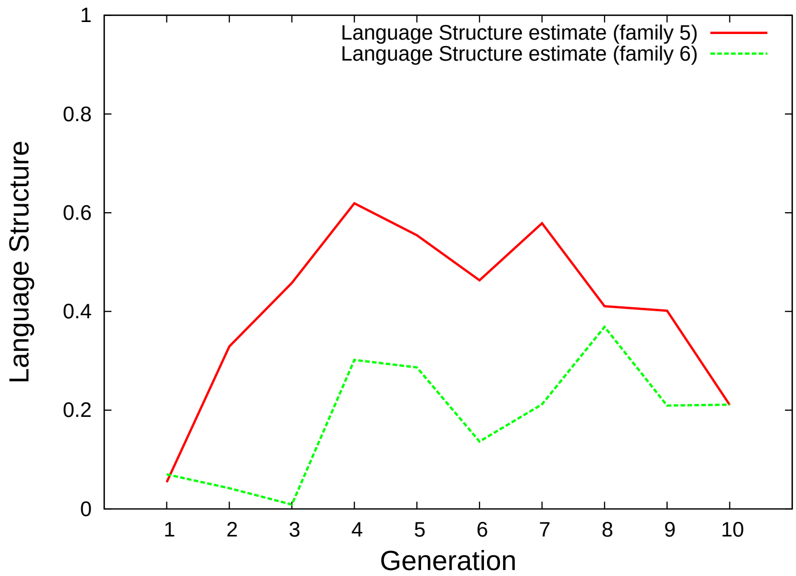


Figure 4.7: Language Structure scores for Families 5 and 6: here we can see that by the end of the experiment, both languages have very little structure in them. There are two points in family 5 however that seem to indicate that there was once structure present.

consistently refer to the motion aspect of the meaning: '-eko' relates to HORIZONTAL, '-kuki' relates to BOUNCE, and '-kiko' relates to SPIRAL. The colour aspect is most clearly marked by the initial letter of the word: 'w-' for BLACK, 'k/h-' for BLUE, and 'p-' for RED. The shape aspect of the meaning-space is less clear-cut, but one way of interpreting the data in table 4.7 is to assume that rather than analyzing the meaning space fully into it's three constituent parts, this subject is making a division of labour such that the first part of the word stands for an individual shape-colour combination. To illustrate this, there is consistent use of the morpheme 'wine-' to refer to a black-circle, 'kun-' meaning blue-circle, 'hun(e)-' meaning blue-square, and so on.

If we examine the data that this subject received from the previous generation (Appendix B.7) we can start to understand how this system may have come about. Although generation 3 had a number of ways of referring to a BOUNCE item (wikiko, kunkuki, ponekuki, winikiko, hunekuki, kunekuki, ponekiko), due to the random sampling of items, all of the items that generation 4 saw had the '-kuki' suffix. In the same way, all of the HORIZONTAL items she saw had the '-eko' suffix. The SPIRAL case was a little different, as here the subject at generation 4 saw: wikiko, hunekiko, kunekuki, wikuki, and poneki. Assuming that she had already decided

	BLACK	BLUE	RED	
HORIZONTAL	wuneko wineko wikeko	huneko kuneko kuneko	puneko poneko poneko	SQUARE CIRCLE TRIANGLE
BOUNCE	winukuki winekuki wikekuki	hunekuki kunkuki kunekuki	punekuki ponekuki ponekuki	SQUARE CIRCLE TRIANGLE
SPIRAL	winekiko winekiko wikiko	hunekiko kunkiko kunekiko	puniko pokiko pokiko	SQUARE CIRCLE TRIANGLE

Table 4.7: Analysis of language at generation 4 of Family 5: this table shows the 'emergent' compositionality that arose in the language. Words are formed by a prefix indicating the shape and colour combination (e.g. kun = blue-circle, wik(e) = black-triangle), and a regular suffix indicating motion (e.g. -kuki = bounce).

that '-kuki' meant BOUNCE, it would make sense for the suffix for SPIRAL to be the one that she saw most often[6]: '-kiko'.

Unfortunately, this 'emergent' compositionality never fully established itself – although a similar system did appear later on at generation 7. This could be due either to the specific properties of the language, or to do with the fact that it unfortunately appeared in the 50% bottleneck condition, which we already know may be less capable of maintaining a compositional system.

4.3.3 Discussion

It was found that filtering out ambiguities that arose over successive generations seemed to prevent either language being learnable to the same extent as witnessed previously. In particular, subjects learning the language in Family 6 appear to have resorted to following a similar strategy to the naive agent. However, two interesting observations were also made. Firstly, the performance of the naive agent actually improved over the course of Family 5, indicating that the language was changing to be more easily learnt by probabilistic strategies. Secondly, two generations within Family 5 seemed on the verge of constructing a compositional system. This leaves us with potentially more questions than answers.

[6] Although there are many factors possibly influencing how salient a particular word is to an individual subject – maybe something about the contrast between -eko, -kuki, and -kiko simply appealed, or maybe it is due to which items she saw first or last during training. All of these suggestions could be investigated further in future work using this empirical framework.

Thinking about what may have caused the increase in performance of the naive agent, it seems likely that a candidate solution could be that the words in the language are becoming more similar to one another. There are several possible reasons for this. Firstly, we must remember that all of the ambiguities are being filtered out in this experiment. This does not actually stop individual participants from creating ambiguities, it just prevents this from being the dominant trend.

Often people will, as a result of imperfect learning or a typo, accidentally introduce a variant that is very close to the word that they were actually aiming for. In an ordinary experiment, this variant will usually be treated as noise if it does not exactly fit the dominant pattern[7]. However, in this run the noise accumulates and becomes significant; we start to see many variants that differ from each other by just one or two letters (e.g. From family 5, generation 10: 'hunekiko', 'hunekike', 'punikike', 'punikiko', 'punike', 'winike', 'wineke' etc.). If this happens extensively throughout the language then the naive strategy will start to do better on the data, as picking any of the seen words should result in a fairly close match to the target word.

The appearance of some kind of 'nearly compositional' system is a little harder to explain. For one thing, it would appear that it could owe its existence largely to chance factors, such as the fact that many words with the same endings were seen. One of the advantages of this framework however is that it allows us to retrospectively analyse interesting scenarios such as these by exactly re-initializing the previous generation's data at the point of interest to see what multiple subjects make of it. If there is an accordance in their collective actions (i.e. we repeatedly see emergent compositionality) then we can conclude that this is more than just a 'one-off'.

[7]For an example of this see Appendix B.6, generations 1, 2, and 3: bouncing-black-triangle ('pokiwe') is seen by generation 2 and incorrectly labelled 'pokinwe'. This item was then seen by generation 3, who promptly corrects the earlier mistake. This was not an isolated example in this family, but was also prevalent when there was established structured ambiguity.

Chapter 5

General Discussion

The aims of this study were initially rather modest; it was hoped that the framework might confirm that language itself adapts to become learnable by its users, and in doing so, might justify the suggestion that it is a useful tool with which to view language evolution, or any other complex culturally transmitted trait learned by observation. This section examines whether it has succeeded in its stated aims, and attempts to look more closely at the issues the results raise, what problems the framework may be usefully applied to, and ways in which the methodology may be improved and extended.

5.1 Summary of Results

The results coming from the three experiments can be summarized as follows:

1. First and foremost, we can confirm the main experimental hypothesis that language evolves to become learnable. This was shown by the fact that languages at later generations were reproduced with more consistency than those at the beginning, in terms of both the seen and the unseen items.
2. The size of the bottleneck may affect the speed at which learners converge on a stable language – a larger bottleneck encourages faster stabilization.
3. The way in which languages evolved to become learnable was by exhibiting *structured ambiguity* – a reduction in the dimensionality of the meaning-space (hence a reduction in the overall number of expressible meanings) occurring in a principled and regular way.
4. A compositional alien language could be initially acquired by subjects.

5. A compositional alien language could only be maintained in the larger bottleneck condition however, as it requires more effort and concentration on behalf of the learner, who benefits from the extra exposure to the system they receive with the larger bottleneck.
6. A successfully maintained compositional system can tolerate some degree of internal change and still remain compositional.
7. Preventing language from expressing structured ambiguity results in the languages no longer becoming easily learnable.
8. In extreme cases (large bottleneck condition) this results in the subjects adopting a naive probabilistic strategy.
9. In less extreme cases (small bottleneck condition) it results in the language subtly changing so as to become easier to learn by probabilistic techniques over time.
10. Two examples of 'emergent' compositionality were observed, but could not be maintained.

It is clear from this summary that the study has proven more successful than it was originally hoped. In particular, the result about the way in which language becomes learnable over time has been a particular surprise. The implications of this finding will be discussed below.

5.2 Ambiguity as Adaptation: Two Routes to Learnability?

Part of the interest in iterated learning stems from a desire to explain why it is that language has the structure it does, hence there is a great deal of focus in the literature into exploring how syntactic features like compositionality emerge. As already discussed, the way that this is explained in an iterated learning account is that it arises as a way of making the languages easier to learn. Although it is understood in an abstract sense that there is more than one way in which language may evolve to become learnable, the fact that compositionality is such a hallmark figure in natural languages means that it very often takes centre stage in any discussion of cultural transmission, learnability, and language evolution.

One of the key findings in this study however, is that compositionality need not be the first port of call en route to learnability. This would not be a surprise if it were to emerge in a model; as previous discussion pointed out, models have to be 'engineered' in order to imitate known biases present in humans concerning a tendency to preserve one-to-one mappings between meanings and signals. The fact that it *did* emerge here, in an experiment involving humans who *already* possess such a bias, is worthy of further investigation.

The first point we should make is that ambiguity is actually rife in natural languages, but relatively absent from formal languages (Wasow et al., 2005; Hoefler, 2006). The reason for this is that ambiguity can lead to misunderstanding, and has often thought to be a major hinderance to successful communication (see earlier quote by De Beule et al. (2006) in 4.1.3). Given this then, why is it that ambiguity is so prevalent? There has been at least one study which attempts to answer this question by looking at the phenomenon within the context of iterated learning. Using a modified ILM Hoefler (2006) shows how syntactic ambiguity arises as an adaptation, both caused and constrained by the transmission bottleneck. In the model, the ambiguity plays a similar role to what we see in the experiment[1] – it increases the learnability of the languages, whilst at the same time, stabilizing them.

By introducing ambiguity that reduced the meaning-space in a predictable and regular way, the language was able to be quickly and faithfully replicated between generations. We can view this another way, with reference to Zipf's Principle of Least Effort (1949). This argument consists of a scenario where there are a finite number of m meanings to be expressed. If a speaker is charged with expressing all of these meanings, then the easiest solution for that speaker is to just utter a single signal for all, as it saves the effort of encoding them individually. If however, there exists a hearer whose job it is to interpret those signals and reassign them to the original meanings, then it would be easier for the hearer if there were m signals to work with, as it saves the effort of having to guess which meaning the speaker is referring to. These two competing forces are engaged in a trade-off, the solution to which lies in a compromise of some number of words less than m but greater than one.

The study suggests that something along these lines is taking place. It is important to realise however, that although the solution the subjects settled on was a positive adaptation in the context of learning this alien language, this would not be a good way of approaching a real language. The strategy employed by the speakers in the experiment is optimal only if there is no requirement to have to successfully decode the signals. This

[1] This is in spite of the fact that the ambiguity Hoefler discusses is syntactic, and the ambiguity witnessed here is lexical

has several implications if we wish to study how language evolves to be learnable via compositionality.

Firstly it suggests that encouraging compositionality to emerge is unlikely to be successful by simply removing any evidence of ambiguity in the input – indeed, this was confirmed empirically during the study. Secondly, it suggests a way in which it might be possible to encourage compositionality to emerge. The learning task as it currently stands means that the subjects are playing a very passive role. The subjects are not really using the language to do anything with, and even if it doesn't make sense to them, they may be reluctant to 'correct' it because they've been told to treat it as an existent language. In order for compositionality to be a 'better' solution than structured ambiguity, the task may have to be changed in such a way that the subject must both learn the strings associated with meanings, and learn the meanings associated with strings. In other words, we need to encourage our subjects to not only be able to produce the strings they have learnt, but also to effectively *use* the language toward some communicative goal that they have a stake in.

5.3 Utility of the Framework

I have illustrated how this framework may be used to explore and confirm findings based on simulation literature that is fairly well known. The fact that it has shown us something new – that there is a potential 'arms-race' between compositionality and structured ambiguity over which will be more successful depending on the task – is a positive sign that more work needs to be undertaken in this area. However, it is impossible to evaluate the utility of a novel methodology such as this without some consideration of its limitations. This section discusses some things that would improve the framework, and also looks at other potential areas within language evolution which may benefit from empirical scrutiny. Finally, it considers some basic extensions to the model.

Limitations

In striving to solve any problem it is often inevitable that new problems, previously inconceived of, will often appear out of nowhere. One such problem concerns the various measures available to interpret the data collected from these studies. The current method by which string similarity is calculated using Levenshtein Distance is not ideal. To understand why, consider the following strings and Levenshtein distances in table 5.1. Here we have an instance of three different strings. Comparing the first two strings, we find we get a highly value, reflecting the fact that there is very little similarity between them. Looking at the second pair of strings however, we find that the Levenshtein distance algorithm assigns the same value as the first pair, in spite of the fact that they are clearly more similar to one another, having been created by swapping around the syllables. This may have adversely affected the scoring of some of Family 6, where this type of syllable substitution occurred frequently.

string one	string two	levenshtein distance
hehima	kuliza	4
hehima	mahehi	4

Table 5.1: Illustration of the Levenshtein Problem: Here we have three words with equal levenshtein distances, despite the fact that the lower pair are simply a rearrangement of one another.

Another problem we have with the framework concerns the number of subjects that are required. Because each *language* is a data-point, and not each participant, this means that in order to perform regular statistical procedures upon the data, we need to have lots of different language families. If you consider that each experiment run so far consists of ten generations (and in fact, as I am about to argue, ten people is actually quite a small number), we are talking about possibly hundreds of people required in order to get enough data to perform statistics on.

This kind of situation is going to require creative solutions – in fact we have already seen one such creative solution in the use of the Monte Carlo Naive Agent, where we could maximize the data we had by comparing it against 1,000 algorithmic responses. A first step in addressing this problem lies in calculating the actual extent of the variation in individual's responses. This could be done by simply taking a random language and, from the same starting point, see where ten different families take it. Isolating the interesting points in the language's evolution can also help direct where appropriate subject resources should be spent. Through careful design and strict controls, the problem of participant numbers is not insurmountable, although it is still an issue to bear in mind.

Applications Within Language Evolution: An Example

Language evolution obviously did not happen overnight. Bickerton (1990) argues that the development of modern human language occurred somewhere between *Homo Erectus* two million years ago, and the appearance of anatomically modern *Homo Sapiens* around 200,000 years ago. In the intervening time between, Bickerton proposes the existence of a protolanguage. This begs the question of what this protolanguage may have looked like. According to Bickerton (2005, p.8), it consisted of:

> "a categorially complete, if severely limited vocabulary of items roughly equivalent to modern words, but lacking a sophisticated phonology and any consistent structure."

This synthetic viewpoint can be contrasted with the analytic view of Wray (2000), who maintains that protolanguage was originally a holistic system, akin to primate calls. The debate gets especially heated when trying to explain how compositionality emerged from these two very different starting points. For Wray, when the number of holophrastic utterances got too large to handle, they start to get analysed into separate constituents, and compositionality is born. For proponents of the synthetic route, the transition from simple words expressing simple concepts to complex arrangements of words expressing complex concepts would have been triggered by some brain development (Bickerton, 2003).

Computational models have been built that provide evidence for both routes (see (Hurford, 2000) for a synthetic account, and (Kirby, 2000) for an analytic version). In a recent article, Bickerton (2005) comments on an example given in Wray (2000) and quite rightly demands that proponents of the analytic view answer the question of how signals could get decomposed if they appear in multiple contradictory contexts. The example that Wray gave is of a language consisting of the following holophrases and their meanings.

The analytic argument is that, given some data like this, there could be chance matches between elements of the signal and elements of the meaning (in this case, the constituent *ma* and some aspect of meaning correlating to *female person* and *beneficiary*) which would lead the hearer to decompose the language. This account is sound, provided that there are no counter-examples. If there are, Wray speculates that one of three types of hypercorrection could occur and prevent a tentative hypothesis from being instantly rejected. However, without any robust data on how people

signal	meaning
tebima	*give that to her*
mupati	*give that to me*
kumapi	*share this with her*
pubatu	*help her*

Table 5.2: Example Language taken from Wray (2000, p.294). Wray claims that given a language like this it would be possible to analytically decompose some overlapping meanings (female + beneficiary) into a partial signal (ma).

react to inconsistent and contradictory data, the argument is rather hard to settle.

Of course, artificial language learning tasks have already been used to investigate the way people interpret inconsistent data. The study by Hudson-Kam and Newport (2005) discussed in the literature review earlier, would maybe suggest that a child would regularize this system, but that an adult would not. However, using a human iterated learning framework may also provide a tool with which to investigate this phenomenon and begin to address Bickerton's query. As mentioned in the discussion in 4.3.3 we can use the framework to set up interesting scenario's, such as the one Wray presents, and see how multiple people interpret the data. We can even go one further than a simple ALL task, and quantify the likelihood of this random chance situation developing in the first place, or examining the future effect that starting off this process of analytic decomposition will have[2].

Extensions to the Framework

Right at the beginning of this study I said that one of the desirable properties of the framework would be to have the ability to explore other modes of cultural transmission from within it. One type of extension to the model would be to do this. The most obvious thing would be to explore both vertical ('parent to child') and horizontal ('peer to peer') transmission at once. There are several different designs that could be chosen.

1. A simple way to incorporate both vertical and horizontal transmission is to perform the experiment in its original form, using the same

[2]Note that if the language had evolved from the random initial conditions into a maintainable compositional language, this would have provided strong support for the analytic route. However, it is also possible to conceive of a method of training that might encourage a synthetic system to emerge. The point however is that we do not know this *a priori* but we possibly now have a means to test it.

alien language as stimuli to a group of participants. At the end of the experiment, all of the subjects responses are collected together and randomly sampled from in order to create the new language.

2. In a slightly more complicated version, two (or more) participants would undergo some degree of training on the same alien language, although in separate rooms. They would be told at the start of the experiment that their task was to learn the alien language, but also, that they must work with an unknown subject in order to co-ordinate their answers to a variety of tests. They will be told ahead of time that one of the tests will involve 'guessing how the alien says X', but they will not actually get to see the alien term for 'X'.
   Instead, at the point where they would usually undergo the first test, subject A will get a chance to select from a set of meaning pictures one that s/he wants B to practise on. These pictures will all be ones that they have not been trained on. Subject A will be able to see the answer that B gives, but will not be able to communicate further. Subject B is then also given an opportunity to select from a set of pictures, and the process repeats for a while. The subjects will then undergo more rounds of training, and practise together until the final test, which as in this study, will consist of both seen and unseen items, some of which will be given as the new input to the next generation of two (or more) participants.

Another way we can extend the experiments is to try to manipulate the task in the way inspired by our discussion of Zipf (1949) earlier. This could be done by simply alternating the testing procedure so that instead of always asking for the description of the picture, it occasionally asked the subject to identify the correct picture when prompted by the description.

It is important to remember that all of these possible extensions exist on top of the fact that there are still many more parameters identified within the modelling literature that have still not been explored. In particular, the structure of the meaning-space and the number of generations would seem likely to be good places to start working on. We have really only seen a snap-shot of all the results that are waiting for us out there.

Chapter 6

Conclusion

This thesis has tried to address the question of why language is structured the way it is, and not some other way. Clearly its current structure has something to do with the way language is acquired, how we as a species have biologically evolved, and crucially, how it is that language is transmitted between speakers. This latter process is the focus of the current paper. It was argued that we learn language by observing the external linguistic actions of others (E-Language) and inducing our own internal representations of it (I-Language). Every time language is transmitted to a new speaker, the process of transforming language from two different forms of representation (E to I) over generations forces language through a learning bottleneck, creating selection pressures that favour structures that can be generalized from, such as compositionality (Hurford, 2000).

So far, most of what we know about this phenomenon (known as iterated learning) comes from computational simulations (Kirby and Hurford, 2002). Whilst this has taught us a great deal about the kinds of learning processes we should expect, what is really required are experimental studies to confirm, and hopefully go beyond the predictions made by the models. However, as yet there is no empirical framework through which to examine a cultural trait as sophisticated and complex as language. A novel methodology was presented here in order to redress this. A subject is initially trained on a subset of a random 'alien' language consisting of 27 pictures (meanings), depicting differently coloured geometric shapes engaged in motion, and strings (signals) that describe them. The subject is then asked to reproduce data they have been trained on, but also, to reproduce data that they have not seen before – thus recreating the learning bottleneck under laboratory conditions.

In order to assess the viability of the framework, a series of six experiments were performed, with ten generations of speakers involved in each.

In the first language family, subjects were trained on an initially random language using a 50% bottleneck (i.e., 50% of the data they were tested on had already been seen during training). The experiment was contrasted with a family trained on a similar language, but with a 75% bottleneck. It was found that the size of the bottleneck was affecting the rate at which languages became stable, but that languages in both instances evolved to become more easily learnt. However, the way that it did so was not by becoming compositional. Instead, the language became ambiguous with respect to specific regions of the meaning space – for instance, by having all bouncing items given the same word, regardless of the other meaning dimensions. In order to reflect the fact that this was not a destructive or random process, this phenomenon was called 'structured ambiguity'.

What the appearance of structured ambiguity does is reduce the size of the meaning-space in a predictable and reliable way, for instance, by consistently not coding in a distinction based on colour. The learners' task is made easier both during word production (as identification of the meanings is now based on fewer interacting variables), and word learning (as the pattern is consistent and there are fewer words to learn). Four further studies were performed in order to investigate whether a compositional system could even be maintained, and also, to explore what happens to languages that are prevented from exhibiting structured ambiguity.

These showed that compositionality could be maintained, and even be robust enough to allow internal change, but only in a system with a larger bottleneck. Additionally, when language was prevented from exhibiting structured ambiguity, it was no longer learnable or stable. There are several conclusions to be drawn from all this. Firstly, the compositional systems were harder to learn than the structured ambiguous ones, although both systems led to language learnability and stability. Secondly, filtering away ambiguity from the input to the learners is maybe not a sufficient pressure in itself to tilt the balance towards favouring a compositional language. For this to occur we may need to develop the framework slightly to make the subjects less passive, and more realistically engaged in a task that requires communication.

Overall, what has been discovered so far is highly promising, and certainly worthy of further analysis and study. Isolated examples of 'emergent' compositionality were found during a run, which in itself is an encouraging sign that this is a framework naturally suited to exploring the kinds of things we are interested in. Furthermore, what we learn about the way in which language is culturally transmitted will have widespread implications for arenas beyond evolutionary linguistics – touching as it does upon possible ways in which *any* cultural artifact is transmitted.

For too long, the field of language evolution has had to content itself with the words "We can only speculate on...", just because we simply have not had the data to go on. Obviously this is not going to provide an answer to every question, but in being able to control cultural transmission in the laboratory, we have gone a step further than modelling such a complex process. We now have a tool that can be used to re-assess much of the previous work that has been undertaken, and a yardstick by which to measure whether we have got an answer that comes anywhere near to the truth revealed by the rigours of experimental verification.

Appendix A

Instructions to Subjects

The following screenshots represent all written instructions to the subjects.

Welcome to Alpha-3-6a in a galaxy far far away. We have encountered an intelligent alien life-form with its own form of language. You must try to learn this language as best you can.

Don't worry if you feel overwhelmed - the alien knows that this is a difficult task for you to master and it will do its best to understand everything that you say.

(Press ENTER to continue)

You will see a series of pictures and the way in which the alien would describe those pictures. Every now and then the alien will test your knowledge of the language by showing you a picture without any description. Simply write what you think the correct response is in the input box provided.

DON'T WORRY IF YOU FEEL YOU HAVE NOT YET MASTERED THE LANGUAGE!

The most important thing is to maintain good relations with the aliens and give it your best shot. ALWAYS GIVE AN ANSWER. That way the aliens will know you are trying. They will go out of their way to try to understand everything you say and they are very patient.

You will be given a break every 5 minutes or so.

If you have any questions please ask the experimentor now.

GOOD LUCK!

(press ENTER to start the tuition)

Appendix B

Language Families

The following tables show the raw data for each generation with a language family. Items in green represent items that were seen by that generation. In order to work out what that generation was trained on, you therefore must look to the left of a word highlighted in green. In cases where the boxes are left blank, this means that the subject did not produce a response for that item.

B.1 Pilot 1

Gen	0	1	2	3	4	5	6	7	8	9	10
	poni	poni	poni	poni	tealawama	tealawama	tealawama		tealawama		
	mowono		lewene		tealawama	tealawama	tealawama		tealawama	tealwama	tealwama
	wogumilo	poni	poni				poni				tealawama
	kekiwe			mahowawa	poni		poni			tealawama	tealawama
	momogona		momowa	mahoha		poni		tealwama			
	puho	momowa	momowaha	tealwama				tealawama			tealawama
	pugune		tealawama		poni	poni		tealawama	tealwama	tealawama	
	luwilika	pimilauke		poni	poni						tealawama
	wikegima	teilauke	tealwama	tealawama			poni	poni	poni	hopa	
	nimewe				lahowama			lewene	lewene		lewene
	miwukuke	hemehe		mahowaha		mahawawa			lewene		
	gewupolu		hemehe		mahowawa	lewene	lewene		lewene		
	kawoha	mewene				lewene	lewene	lewene		lewene	
	gili										
	weni		mewene		mahowawa	mahowawa		lewene		lewene	lewene
	gage	meheha	meheha			lewene	lewene	lewene			lewene
	hemeha	meheha			mahowawa	lewene			lewene	lewene	
	pigu	puoie	meheha			lewene		lewene		lewene	lewene
	popena	mowama	lewene	tealwama	mehehe	tealwama	tohowama	tohowama	wohawama		
	kilikope					tealwama	wohawama	wohawama		tealwama	tealwama
	giwe		pilikaue	mehehe			tealwama			tealwama	tealwama
	gakawu	pilkaue		lewene	lewene				tealwama		tealwama
	kigegake			lewene		mehehe		tealwama	tohawama	tealawama	
	weku	gemoma		memeh	mehehe					tealwama	tealwama
	hiku				tealwama		tohowama	tealawama	tealawama	hopiwama	
	wekuke	milikaue		mehehe			tealwama	wohawama	tealawama	tealwama	
	pehakigo		milikaue	mehehe	mehehe		tealwama		tealwama		tealwama

B.2 Pilot 2

Gen	0	1	2	3	4	5	6	7	8	9	10
	nakegama	nakegama	kini	kepola	kipola	kepola	kepola	kepali	kepali	kepoli	kepoli
	kiki	kini	kihoni	kipola	kepola	kepoli	kepola	kepali	kepali	kepoli	kepoli
	kipi	kiki	kiki	kepola	kepola	kepola	kepola	kepali	kepali	kepoli	kepola
	nakenihi	kini	ninama	kepola	kahini	kipola	kepali	kepali	kepoli	kepoli	kepoli
	hohiki	gohina	kihoni	gilina	kipola	kipola	kepoli	kepali	kepali	kepoli	namini
	nikipi	kehoni	nanimi	kini	kipoli	kepoli	kepoli	kepali	kepali	kepoli	kepoli
	nahiga	piholic	kini	kini	kipoli	kepoli	kepoli	kepali	kepali	kepoli	kepoli
	kepi	kinigama	nihoni	kini	kipola	kepoli	nanimi	kipola	kepali	namini	namini
	gake	nakegama	kiki	kiki	kipola	kepoli	kepali	kepali	kepali	kepoli	kepoli
	hinikena	kinali	kipolic	nanimi	nanima	nanima	nanima	nanimi	nanimi	namini	namini
	nahima	namini	kihoni	nanimi	nanima	nanima	nanima	nanimi	nanimi	namini	namini
	ninahi	nimali	kihoni	nanoma	nanima	nanima	nanima	nanima	nanima	namini	namini
	gahoni	golini	kipolic	nanima	nanimi	nanimi	nanimi	nanimi	nanimi	namini	namini
	naniga	hinima	nanimi	nanima	nanimi	nanimi	nanimi	nanimi	nanimi	namini	namini
	himahohi	nanima	nanimi	nanima	nanimi	nanimi	nanimi	nanima	nanima	namini	namini
	kenihi	pihoni	kepoli	nanimi	nanimi	nanimi	nanimi	nanimi	nanimi	namini	namini
	pihoma	gohili	kipolic	nanima	nanima	nanimi	nanimi	nanimi	nanimi	namini	namini
	hikiga	pihoni	kihoni	nanima	nanimi	nanimi	nanimi	nanima	nanima	namini	namini
	kini	pikoli	kini	nahimi	kepola	kepola	kepola	kepola	kepola	kepola	namini
	nimaki	kinipi	gihoni	kepola	kipola	kepoli	kepola	kipola	kipola	namini	namini
	nihi	nanima	namini	kepola	kepola	kepola	kepola	kepola	kepola	kepola	kepola
	nakikema	nakoli	kihoni	nanima	kipola	kepoli	kepoli	kepola	kepola	kepola	namini
	kinipi	kehopi	kipolic	kahini	kahini	nanimi	kepoli	kipola	kipola	namini	namini
	honapini	nanili	kiholi	kepoli	kipola	kipola	kepola	kepola	kepola	kepola	kepola
	kehoma	kohipi	nanima	kahimi	kepoli	kepoli	kepola	kepola	kepola	kepola	namini
	hokehopi	piholi	nanima	kahini	kipol	kepoli	kipola	kipola	kipola	kipoli	namini
	gahipihi	gomali	gilina	kepoli	kipola	kipola	kepola	kepola	kepola	kepola	kepola

B.3 Family 1

Gen	0	1	2	3	4	5	6	7	8	9	10
	kinimapi	kimei	hepini	tuge	tupim	tupim	minku	miniku	miniku	miniku	miniku
	wikuki	miwn	miniku	tupim	tupim	tupim	miniku	miniku	miniku	miniku	miniku
	kikumi	miheniw	hepini	tupim	tupim	tupim	miniku	miniku	miniku	miniku	miniku
	miwiniku	pemini	nige	miniku	mihunu	mihunu	miniku	miniku	tupim	tupim	tupim
	pinipi	kupini	tuge	tuge	tupim	tupim	tupin	miniku	tupim	tupim	tupim
	kihemiwi	pon	mihenu	mihunu	miniku	tupim	tupim	tupim	tupim	tupim	tupim
	miwimi	poi	poi	poi	poi	miniku	miniku	miniku	tupin	tupin	tupin
	nipi	mhip	mpo	tuge	miniku	tupim	tupin	tupin	tupin	tupin	tupin
	wige	kuwpi	tupim	miniku	miniku	miniku	tupin	miniku	tupin	tupin	tupin
	nihepi	mip	nige	tuge	tuge	tuge	tuge	tuge	tuge	tuge	tuge
	wigemi	mpo	nige	tuge	tuge	tuge	tuge	tuge	tuge	tuge	tuge
	mahekuki	miniku	tuge	tuge	tuge	tuge	tuge	tuge	tuge	tuge	tuge
	wimaku	nige	nige	mihenu	tuge	tuge	tuge	tuge	tuge	tuge	tuge
	miniki	miniku	tuge	tuge	tuge	tuge	tuge	tuge	tuge	tuge	tuge
	gepinini	poh	tuge	tuge	tuge	tuge	tuge	tuge	tuge	tuge	tuge
	wikima	tuge	nige	nige		tuge	tuge	tuge	tuge	tuge	tuge
	nipikuge	tuge	tuge	tuge	tuge	tuge	tuge	tuge	tuge	tuge	tuge
	hema	weg	mpo	tuge	tuge	tuge	tuge	tuge	tuge	tuge	tuge
	pikuhemi	kuhepi	hepini	tupim	tupim	tupim	poi	poi	poi	poi	poi
	kimaki	wige	tupim	tupim	tupim	tupim	poi	poi	poi	poi	poi
	pimikihe	mie	tupim	tupim	tupim	tupim	poi	poi	poi	poi	poi
	gepihemi	hepinimi	hepini	mihenu	tupim	tupim	poi	poi	poi	poi	poi
	kunige	himini	miniku	tupim	tupim	tupim	tupin	poi	poi	poi	poi
	miki	hipe	tupim	tupim	tupim	tupim	tupim	tupim	poi	poi	poi
	mihe	pobo	nige	poi	poi	poi	poi	poi	poi	poi	poi
	winige	tupim	tupim	tupim	tupim	tupim	tupin	tupin	poi	poi	poi
	kinimage	hipe	poi	tupim	tupim	tupim	tupim	poi	poi	poi	poi

B.4 Family 2

Gen	0	1	2	3	4	5	6	7	8	9	10
	kawi	minake	gokame	gomina	gomina	gomina	gomina	goina	goina	goina	gomia
	gopunena	gomehoki	gokame	gomare	gomina	gomina	gomina	gomina	gomina	gomina	gomina
	hogona	mihara	gokame	gomina	gomina	gomina	gomina	gomina	gomina	gomina	gomina
	pukakape	honare	honare	honare	honare	honare	honare	honare	horare	horare	horare
	honemi	honare	honare	honare	honare	honare	honare	honare	honare	honare	honare
	gopepuho	gopuka	honare	honare	honare	honare	honare	honare	honare	honare	honare
	hokapumi	kakawiha	kakawi	kakawi	kakawi	kakawi	kakawi	kakawi	kakawi	kakawa	kakawa
	nemiho	kakawihe	kakawi	kakawi	kakawi	kakawi	kakawi	kakawa	kakawa	kakawa	kakawa
	kakawiho	kakawiha	kakawi	kakawi	kakawi	kakawi	kakawi	kakawa	kakawa	kakawa	kakawa
	gokami	gokare	gokami	gomia	gomia	gomia	gomia	gomia	gomia	gomia	gomia
	mina	gokami	mina	gomina	gomia	gomia	gomia	honare	honare	horare	horare
	gomipe	mina	mina	gomina	gomia	gomia	gomia	gomia	gomia	gomia	gomina
	goho	goho	goho	goho	honare	honare	honare	honare	honare	horare	horare
	mipeho	goho	goho	goho	honare	honare	honare	honare	honare	honare	honare
	pumi	goho	goho	goho	honare	honare	honare	honare	honare	honare	honare
	napu	puwa	puwa	puwa	puwa	puwa	puwa	kakawa	kakawa	kakawa	kakawa
	pena	puwa	puwa	puwa	kakawa	kakawa	kakawa	kakawa	kakawa	kakawa	kakawa
	wipune	wikame	puwa	puwa	kakawa	kakawa	kakawa	kakawa	kakawa	kakawa	kakawa
	pehomi	gokami	minare	minare	gomina	goina	goina	gomina	goina	gomia	gomia
	mine	minare	minare	minare	gomina	gomina	gomina	gomina	gomina	gomina	gomina
	pumihopu	hokikapa	minare	minare	gomina	gomina	gomina	gomina	gomina	gomina	gomina
	miwinana	honaka	honaka	horare	honare	honare	horare	horare	horare	horare	horare
	honaka	honaka	honake	horare	honare	horare	horare	horare	honare	honare	honare
	negona	honaka	honake	horare	horare	horare	horare	horare	honare	honare	honare
	naka	kakawi	kakawi	kakawa	kakawa	kakawa	kakawa	kakawa	kakawa	kakawa	kakawa
	howimi	mepumi	kakawi	kakawi	kakawa	kakawa	kakawa	kakawa	kakawa	kakawa	kakawa
	kahogo	wikame	kakawi	kakawa	kakawa	kakawa	kakawa	kakawa	kakawa	kakawa	kakawa

B.5 Family 3

Gen	0	1	2	3	4	5	6	7	8	9	10
●	hekiha	hekoha	hekoha	hekohi	magohe	megohe	megohe	magohe	megome	megoma	megome
●	hekimu	hekomu	hekomu	hekomu	magomu	megomu	megomu	megohe	megohe	megome	megome
●	hekipo	hekopo	hekopo	hehopo	magopo	wikipo	wikipo	wikipo	magame	haheke	haheto
■	hegoha	hegoha	hegoha	hekoha	magohe	megohe	megohe	megohe	wikipu	megohe	magomo
■	hepimu	hegomu	hegomu	hegoho	magomu	megomu	megomu	megomu	megohe	megohe	magohe
■	hegopo	hegopo	hegopo	hegipo	magopo	wikipo	wikipo	wikipo	heheke	hehake	haheto
▲	hehiha	hehiha	hehoha	hehohi	megohe	magehe	magohe	magohe	magahe	megohe	megomo
▲	hehimu	hehimu	hehomu	hehamu	hegomu	magemu	hahemu	magahe	megohe	megohe	megohe
▲	hehipo	hehipo	hehopo	hehipo	hagopo	wikipo	wikipo	wikipo	wikipo	wikipu	wikipo
●	wikiha	wikiha	wikoha	wikohi	magohe	magohe	magohe	magahu	haheke	magohe	megopu
●	wikimu	wikipo	wikomu	wikomu	makimu	magomu	megomu	magohe	magame	magome	megopu
●	wikipo	wikipo	wikipo	wikopo	makipo	wikipo	wikipo	wikipo	heheke	hakeko	hahepu
■	wigoha	wigiha	wigoha	wihohi	makihe	magohe	megohe	haheku	megemo	wikipu	magome
■	wigomu	wigimu	wigomu	wikohe	hahemu	hahemu	megomu	megohe	magame	magepo	magopu
■	wigopo	wigipo	wigipo	wigopo	wikipo	wikipo	wikipo	wikipo	heheke	haheke	hahepu
▲	wihiha	wikiha	wihiha	wihihe	hahihe	magohe	magohe	magahe	wikipo	magoma	megoho
▲	wihimu	wihimu	wihiha	wihimu	hahimu	hahemu	hahemu	magahu	megohe	magome	megoho
▲	wihipo	wihipo	wihopo	wihipo	hakipo	wikipo	wikipo	wikipo	wikipo	haheke	wikipu
●	makiha	makiha	magoma	magohe	wihahe	hahehe	megohe	megohu	magahe	megome	magopu
●	makimu	makimu	makomu	makomu	mahimu	magomu	megomu	megohe	magame	megohe	megopu
●	makipo	makipo	magopo	magopo	wikipo	wikipo	wikipo	wikipo	heheke	haheke	haheki
■	magoha	magoha	magoma	makohi	makihe	magohe	magahe	magohe	magahe	megome	magoho
■	mapimu	magimu	magomu	mahimu	mahomu	hahemu	megomu	magahe	magame	magome	magohe
■	magopo	magipo	magipo	mahopo	hahopo	wikipo	wikipo	wikipo	wikipo	wikipo	hahepu
▲	mahiha	mahiha	mahoha	mahiho	hegohe	hahehe	magohe	megohe	wikipo	magope	megoho
▲	mahimu	mahimu	mahimu	mahimu	hegomu	hahemu	hahemu	megohe	megohe	megome	megohe
▲	mahipo	mahipo	mahopo	mahipo	wikipo	wikipo	wikipo	wikipo	wikipo	wikipo	wikipo

B.6 Family 4

Gen	0	1	2	3	4	5	6	7	8	9	10
	ponowe	ponowe	ponowe	ponowe	ponowe	ponowe	ponowe	ponowe	lusawe	posawe	posawe
	pononi	pononi	pononi	ponoki	pononi	pononi	pononi	pononi	posane	posane	posane
	ponona	ponona	ponona	ponona	ponona	ponona	ponona	ponona	posona	posana	posana
	pomewe	pomewe	pomewe	pomewe	pomewe	pomewe	pomiwe	pomowe	pomawe	pomowe	pomawe
	pomeni	pomeni	pomeni	pomeni	pomeni	pomeni	pokiwe	pononi	pomane	pomane	pomane
	pomena	pomena	pomena	pomena	pomena	pomena	pomina	pomona	pomana	pomana	pomana
	pokiwe	pokiwe	pokinwe	pokiwe	pokiwe	pokiwe	pokiwe	ponowe	pokowe	pokowe	pokawe
	pokini	pokini	pokini	pokini	pokini	pokini	pokini	pokoni	pokane	pokane	pokane
	pokina	pokina	pokina	pokini	pokina	pokina	pokina	pokena	pokana	pokana	pokana
	kanowe	kanowe	kanowe	kanowe	kanowe	kanowe	kinowe	kusawa	kasawe	kasowe	kasawe
	kanoni	kanoni	kanoni	kanoni	kanoni	kanoni	kanoni	kisene	kasane	kasane	kasane
	kanona	kanona	kanona	kanona	kanona	kanona	kanona	kasona	kanana	kasana	kasana
	kamewe	kamewe	kamenwe	kamewe	kamewe	kamewe	kakawe	kakawa	kamowe	kamowe	kamawe
	kameni	kameni	kameni	kameni	kameni	kameni	kamini	kamoni	kamane	kamane	kamane
	kanona	kamena	kamena	kanona	kanona	kamena	kunona	kisona	kamana	kamana	kamana
	kakiwe	kakiwe	kakina	kakina	kakiwe	kakiwe	kukiwe	kakawe	kakawe	kakawe	kakawe
	kakini	kakini	kakina	kakina	kakini	kakini	kakine	kukene	kakane	kakane	kakane
	kakina	kakima	kakina	kakina	kakini	kakina	kokina	kasana	kakana	kakana	kakana
	lunowe	lunona	lunowe	lunowe	lunowe	lunowe	lunowe	lunowe	lusawe	lusawe	kasawe
	lunoni	lunoni	lunoni	lunoni	lunoni	lunoni	lunoni	lunone	lusane	lusane	kasane
	lunona	lunona	lunona	lunona	lunona	lunona	lunona	lunona	lusana	lusana	lasana
	lumewe	lumewe	lumewe	lumewe	lumewe	lumewe	lukiwe	lumowe	lumawe	lumowe	lamawe
	lumeni	lumeni	lumeni	lumeni	lumeni	lumeni	lukoni	lunoni	lumone	lumane	lamane
	lumena	lumena	lumena	lumena	lumena	lumena	lumena	lumoni	kamana	lumana	lamana
	lukiwe	lukiwe	lukinwe	lukiwe	luniwe	lukiwe	lukewe	lukowe	lukawe	lukawe	lakawe
	lukini	lukini	lukini	lukini	lukini	lukini	lukoni	lukoni	lukane	lukane	lakane
	lukina	lukina	lukina	lukina	lukina	lukina	lukana	lukana	lukana	lukana	lakana

B.7 Family 5

Gen	0	1	2	3	4	5	6	7	8	9	10
●	huhunigu	pikoku	wikiko	wikiko	winekuki	winekuki	winikike	winikiko	wunkiko	winikiko	winikiko
●	kemuniwa	huniki	hukiki	kunkuki	kunkuki	hunekuki	honekiko	honekiko	kunkike	hunekiko	hunekiko
●	kihupo	piko	pokiko	ponekuki	ponekuki	ponekuki	punekiko	ponekiko	punkiko	punekiko	punekiko
■	wakiki	wukiki	winekiko	winikiko	winukuki	winikuki	winikeko	winikiko	winikiko	winikiko	winikiko
■	pokikehu	ponuko	kunikeko	hunekuki	hunekuki	hunekuki	kunekiko	kunekiko	ponekiko	hunekiko	hunekiko
■	waguhuki	poku	ponekiko	ponekuki	punekuki	punikuki	ponekiko	punekiko	pinkiko	punkiko	hunekiko
▲	nihu	kikiki	kikiki	winekiko	wikekuki	wanikuki	winikiko	winikiko	winekiko	winekiko	punikiko
▲	niguki	hukeko	hukiki	kunekuki	kunekuki	kunikuki	kunekiko	kikekiko	pinekiko	ponikiko	winikike
▲	koni	koni	ponekiko	ponekiko	ponekuki	punekuki	punekiko	punikiko	punkiko	punkiko	punkiko
●	muwapo	wuniki	wineko	wineko	wineko	wineke	wineke	wineke	wuneke	winekike	punike
●	powa	pinokiki	huneko	kuneko	kuneko	kunike	honeke	honeke	kineke	hunike	wineke
●	hukinimu	kuniko	ponukeko	poneko	poneko	ponike	punike	ponike	puneke	punike	winikike
■	wako	wako	wikeko	wineko	wuneko	wanike	wineke	winike	wineke	winike	puneke
■	hukeko	ponikio	huniko	huneko	huneko	hunike	kuneke	kuneke	huneke	ponike	hunekike
■	pohumu	hukeko	ponekuko	poneko	puneko	punike	punike	puneke	puneke	ponike	punike
▲	muko	wakiki	kineko	wineki	wikeko	wineke	wineke	winike	wunike	winike	wineke
▲	kokeguke	piniko	kuneko	kuneko	kuneko	hunike	kuneke	punike	honike	huneke	wineke
▲	kimu	koniki	pokiko	poneko	poneko	punike	punikiko	punike	punike	ponike	wineke
●	kekewa	wiki	wiki	wikiko	winekiko	winikike	winikeke	winekike	winikike	winike	winekike
●	komuhuke	ponukiko	huki	kunekuki	kunkiko	kunikike	honekiko	honekike	kinike	ponike	hunikike
●	kopo	ponikiko	poniki	poneko	pokiko	punikike	punekike	ponikike	poneike	ponike	punikiki
■	huwa	ponikiko	wineko	wikuki	winekiko	winekiko	winikike	winekike	winike	winikike	winikike
■	hukike	hukeke	hunekiki	hunekiko	hunekiko	kunike	kunekike	kunekike	kinkike	hunike	punkike
■	ponikiko	ponikiko	ponekuki	poneki	puniko	punekiko	punikike	punekike	punkike	ponike	punikike
▲	kowagu	winiko	wineki	winuki	wikiko	wanikike	winikike	winkike	winkeke	winike	winikike
▲	kokihuko	hukiki	hunekiko	hunekiko	kunekiko	kunekike	kinekike	kinekike	honkike	huneke	winike
▲	kiwanike	kuniko	pokiko	ponekuki	pokiko	punikike	ponekiko	ponekiko	pinkike	punike	winikike

B.8 Family 6

Gen	0	1	2	3	4	5	6	7	8	9	10
●	pawamu	gopumi	wima	gepoko	pygamo	gemugo	mygempo	vimugo	vimugame	vinegame	pogem
●	gopeko	peko	pakoke	gemegame	winugame	pogame	mypemgo	penegame	penegame	pinegame	pogame
●	muke	pakoke	wima	gepemy	gymugo	myugo	pemugame	pegame	pegame	pagame	pinegame
■	gokokoni	gopawe	wimegamu	gegogo	gemegame	penegame	penegame	vimugo	vimugame	vinegame	vinegame
■	pewakeni	pumewiwa	penigamu	gegumy	gemegame	pegame	pygempo	gepgo	penegame	vinegame	vinegame
■	pakoke	gopoke	kamigamu	wamugamy	gemegame	pegame	pemugo	pegepgo	pegepgo	pegom	vinegame
▲	kope	wunigamu	wunigamu	winugame	winugame	winugame	winugame	vimugame	vimogame	vimegame	pinegame
▲	wako	wamu	wamu	pegamy	winugame	winugame	winugo	vimugo	penegame	gempo	gempo
▲	wikewape	pewawe	pakoke	winugame	penegame	penegame	mypego	vimugo	pegame	pinegame	vimegame
●	niwapake	wima	pago	pagumy	gemgo	gemgo	gego	mypego	mipego	mygamma	gemco
●	keko	pawikako	pagamu	gego	gego	gego	pemugo	migypgo	mypeg	gemco	gemco
●	niwa	pawame	gopawe	gego	pape	gemgo	pemgy	pego	pe	miopy	gemco
■	gowi	wingumi	kego	keko	gego	pape	pape	migypgo	mipeggo	pegym	pogem
■	pawani	pewu	pakoke	gupumy	gopoko	pape	pemgo	gepgo	penego	pegym	pegem
■	gopawa	pamuwawi	pape	pape	pape	pape	gemgo	pegypgo	pego	miopy	vimugo
▲	pamu	wima	gopumi	gupogp	pape	gego	memgo	mygo	vimugo	pe	pogem
▲	panimu	pagake	gopumi	gego	pape	gemgo	mepegy	gemgo	gemgo	pe	pegem
▲	pamugo	pumago	pumago	gepego	pape	gemgo	mypego	pe	pe	pe	pe
●	muniwiwa	kewi	kewi	penegame	gepego	mipegy	mypegy	mygemgo	vimugo	vimugo	vimugo
●	kewi	piwawe	pagamu	gukepo	gepego	mipeg	mipego	migepgo	mypeggo	mipogy	pogem
●	pani	pumikako	panigamu	pogumy	mipego	mipego	mipego	mipego	mipego	migemco	vimugo
■	wimu	kewi	gopoke	keko	gepmy	pegamy	gepmy	migemgo	mypeggo	mygamma	vimugo
■	mukepa	pape	kegami	kegopo	pogomy	gepmy	pemgy	gepgo	penego	miogy	vimugo
■	kepewipa	winipawe	wimigamu	gewugo	pegamy	gepamy	pemygo	pegepgo	mipeggo	miopy	migemco
▲	munikego	kewi	kewi	gepogame	mepegy	mypegy	mypegy	mygepgo	vimugo	miopy	miope
▲	pape	kewi	winugami	penegame	gopomy	mipegy	mipegy	gemgo	penego	miope	miopy
▲	niwipa	pumi	wigamu	gemugo	pogumy	pogomy	mypegy	mygepgo	mipego	pygypo	miope

References

Aunger, R., editor (2000). *Darwinizing Culture: The Status of Memetics as a Science*. Oxford: Oxford University Press.

Bangerter, A. (2000). Transformation between scientific and social representations of conception: The method of serial reproduction. *British Journal of Social Psychology*, 39:521–535.

Bartlett, F. (1932). *Remembering: A Study in Experimental and Social Psychology*. Cambridge University Press.

Batali, J. (1998). Computational simulations of the emergence of grammar. In Hurford, J., Steddert-Kennedy, M., and Knight, C., editors, *Approaches to the Evolution of Language: Social and Cognitive Bases*, pages 404–426. Cambridge: Cambridge University Press.

Batali, J. (2002). The negotiation and acquisition of recursive grammars as a result of competition among exemplars. In Briscoe, E., editor, *Linguistic Evolution through language acquisition: formal and computational models*, pages 111–172. Cambridge: Cambridge University Press.

Bates, E., Thal, D., Finlay, B., and Clancy, B. (2003). Early language development and its neural correlates. In Rapin, I. and Segalowitz, S., editors, *Child Neurology*, volume 6 of *Handbook of Neuropsychology*. Amsterdam: Elsevier, 2 edition.

Bickerton, D. (1981). *The Roots of Language*. Ann Arbor, WI: Karoma.

Bickerton, D. (1990). *Language and Species*. Chicago: University of Chicago Press.

Bickerton, D. (2003). Symbol and structure: A comprehensive framework for language evolution. In Christiansen, M. and Kirby, S., editors, *Language Evolution*, pages 77–93. Oxford: Oxford University Press.

Bickerton, D. (2005). Language evolution: A brief guide for linguists. *Lingua, (In Press)*.

Biro, D., Inoue-Nakamura, N., Tonooka, R., Yamakoshi, G., Sousa, C., and Matsuzawa, T. (2003). Cultural innovation and transmission of tool use in wild chimpanzees: evidence from field experiments. *Animal Cognition*, 6:213–223.

Brighton, H. (2002). Compositional syntax from cultural transmission. *Artificial Life*, 8(1).

Caldwell, C. and Whiten, A. (2004). Testing for social learning and imitation in common marmosets, callithrix jacchus, using an artificial fruit. *Animal Cognition*, 7:77–85.

Carruthers, P. and Boucher, J. (1998). *Language and Thought: Interdisciplinary Themes*, chapter Introduction: Opening up Options, pages 1–18. Cambridge University Press.

Cavalli-Sforza, L. and Feldman, M. (1981). *Cultural transmission and Evolution: A Quantitative Approach*. Princeton University Press.

Chomsky, N. (1986). *Knowledge of Language*. Praeger.

Christiansen, M. (2000). Using artificial language learning to study language evolution: Exploring the emergence of word order universals. In Desalles, J. and Ghadakpour, L., editors, *The Evolution of Language: 3rd International Conference*, pages 45–48. Ecole Nationale Superieure des Telecommunications, Paris.

Christiansen, M. and Devlin, J. (1997). Recursive inconsistencies are hard to learn: A connectionist perspective on universal word order correlations. In *Proceedings of the 19th Anuual Cognitive Science Society Conference*, pages 113–118. Mahwah, NJ: Laurence Erlbaum Associates.

Christiansen, M. and Ellefson, M. (2002). Linguistic adaptation without linguistic constraints: The role of sequential learning in language evolution. In Wray, A., editor, *Transitions to Language*, pages 335–358. Oxford: Oxford University Press.

Christiansen, M. and Kirby, S. (2003a). Language evolution: Consensus and controversies. *Trends in Cognitive Science*, 7(7):300–307.

Christiansen, M. and Kirby, S. (2003b). Language evolution: The hardest problem in science? In Christiansen, M. and Kirby, S., editors, *Language Evolution*, pages 1–16. Oxford: Oxford University Press.

Cooper, R. (2002). *Modelling High-Level Cognitive Processes*. Laurence Erlbaum Associates.

Cornish, H. (2005). The role of meaning within the iterated learning model of language evolution. Unpublished, The University of Edinburgh.

Custance, D., Whiten, A., Sambrook, T., and Galdikas, B. (2001). Testing for social learning in the "artificial fruit" processing of wildborn oragutans

(*Pongo pygmaeus*), Tanjung Putin, Indonesia. *Animal Cognition*, 4:305–313.

Darwin, C. (1859). *The Origin of Species*. London: Penguin, 1968.

Dawkins, R. (1976). *The Selfish Gene*. Oxford: Oxford University Press, 30th anniversary edition.

De Beule, J., De Vylder, B., and Belpaeme, T. (2006). A cross-situational learning algorithm for damping homonymy in the guessing game. In Rocha, L., editor, *Artificial Life X*, pages 466–472. Cambridge, MA: MIT Press.

Deacon, T. (1997). *The Symbolic Species*. London: Penguin.

Dennett, D. (1995). *Darwin's Dangerous Idea*. London: Penguin.

Ellefson, M. and Christiansen, M. (2000). Subjacency constraints without universal grammar: Evidence from artificial language learning and connectionist modeling. In *The Proceedings of the 22nd Annual Conference of the Cognitive Science Society*, pages 645–650. Mahwah, NJ: Lawrence Erlbaum.

Fay, N., Garrod, S., MacLeod, T., Lee, J., and Oberlander, J. (2004). Design, adaptation and convention: The emergence of higher order graphical representations. In *Proceedings of the 26th Annual Conference of the Cognitive Science Society*, pages 411–416.

Fitch, T. and Hauser, M. (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science*, 303:377–380.

Fox, A. (1995). *Linguistic Reconstruction: An Introduction to Theory and Method*. Oxford: Oxford University Press.

Francis, N. and Kucera, H. (1982). *Frequency Analysis of English Usage: Lexicon and Grammar*. Boston: Hoghton Mifflin.

Galantucci, B. (2005). An experimental study of the emergence of human communication systems. *Cognitive Science*, 29:737–767.

Hare, M. and Elman, J. (1994). Learning and morphological change. *Cognition*, 56:61–98.

Hauser, M. D. and Fitch, W. T. (2003). What are the uniquely human components of the language faculty? In Christiansen, M. and Kirby, S., editors, *Language Evolution*, pages 158–182. Oxford University Press.

Hoefler, S. (2006). Why has ambiguous syntax emerged? In Cangelosi, A., Smith, A., and Smith, K., editors, *The Evolution of Language: Proceedings of the 6th International Conference (EVOLANG'06), Rome, Italy*, pages 123–130. World Scientific.

Horner, V., Whiten, A., and de Waal, F. (In Press). Faithful replication of foraging techniques along cultural transmission chains by chimpanzees and children. *PNAS*.

Hudson-Kam, C. and Newport, E. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development*, 1(2):151–195.

Hurford, J. (2000). Social transmission favours linguistic generalisation. pages 324–352. Cambridge: Cambridge University Press.

Hurford, J. (2002). Expression/induction models of language evolution: dimensions and issues. In Briscoe, E., editor, *Linguistic Evolution through Language Acquisition: Formal and Computational Models*. Cambridge: Cambridge University Press.

Hurford, J. (2005). Computer modelling widens the focus of language study. In Tallerman, M., editor, *Language Origins: Perspectives on Evolution*. Oxford: Oxford University Press.

Jablonka, E. and Lamb, M. (2005). *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioural, and Symbolic Variation in the History of Life*. Cambrisge, MA: MIT Press.

Kashima, Y. (2000). Maintaining cultural stereotypes in the serial reproduction of narratives. *Personality and Social Psychology Bulletin*, 26:594–604.

Kegl, J. (1994). The nicaraguan sign language project: An overview. *Signpost*, 7:24–31.

Kirby, S. (2000). Syntax without natural selection: How compositionality emerges from vocabulary in a population of learners. In Knight, C., editor, *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*, pages 303–323. Cambridge University Press.

Kirby, S. (2001). 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.

Kirby, S. (2002a). Learning, bottlenecks and the evolution of recursive syntax. In Briscoe, E., editor, *Linguistic Evolution Through Language Acquisition: Formal and Computational Models*. Cambridge University Press.

Kirby, S. (2002b). Natural language from artificial life. *Artificial Life*, 8(2):185–215.

Kirby, S. and Hurford, J. (2002). The emergence of linguistic structure: An overview of the iterated learning model. In Cangelosi, A. and Parisi, D., editors, *Simulating the Evolution of Language*, pages 121–148. Springer-Verlag.

Levenshtein, V. (1966). Binary codes capable of correcting deletions, insertions and reversals. *Soviet Physics Doklady*, 10(8):707–710.

Lyons, J. (1995). *Linguistic Semantics*. Oxford: Oxford University Press.

Markman, E. and Wachtel, G. (1988). Children's use of mutual exclusivity to constrain the meaning of words. *Cognitive Psychology*, 20:121–157.

Mesoudi, A., Whiten, A., and Dunbar, R. (2006a). A bias for social information in human cultural transmission. *British Psychological Society*, 97:405–423.

Mesoudi, A., Whiten, A., and Laland, K. (2006b). Towards a unified science of cultural evolution. *Behavioral and Brain Sciences*, 29:329–383.

Pinkal, M. (1995). *Logic and Lexicon*. Dordrecht: D. Reidel.

Pinker, S., editor (1984). *Language Learnability and Language Development*. Cambridge, MA: Harvard University Press.

Rendell, L. and Whitehead, H. (2005). Culture in whales and dolphins. *Behavioral and Brain Sciences*, 24:309–382.

Richerson, P. and Boyd, R. (2005). *Not By Genes Alone: How culture transformed human evolution*. Chicago: University of Chicago Press.

Saffran, J. (2001). The use of predictive dependencies in language learning. *Journal of Memory and Language*, 44:493–515.

Saffran, J., Aslin, R., and Newport, E. (1996). Statistical learning by 8-month olds. *Science*, 274:1926–1928.

Skinner, B. (1957). *Verbal Behavior*. New York: Appleton-Century-Crofts.

Slobin, D. (1977). Language chainge in childhood and history. In Macnamara, J., editor, *Language Learning and Thought*, pages 185–221. London: Academic Press.

Smith, K. (2003). Learning biases and language evolution. In Kirby, S., editor, *Proceedings of Language Evolution and Computation, Workshop/Course at ESSLLI*, pages 22–31. Vienna.

Smith, K. and Hurford, J. (2003). Language evolution in populations: extending the iterated learning model. In Banzhaf, W., Christaller, T., Dittrich, J., Kim, J., and Ziegler, J., editors, *Advances in Artificial Life - Proceedings of the 7th European Conference on Artificial Life (ECAL), Lecture Notes on Artificial Intelligence*, volume 2801, pages 507–516. Springer-Verlag.

Steels, L. (2006). Experiments on the emergence of human communication. *Trends in Cognitive Sciences*, 10(8):347–349.

Teal, T. and Taylor, C. (2000). Effects of compression on language evolution. *Artificial Life*, 6(2).

Thomason, S. (2001). *Language Contact: an introduction*. Washington: Georgetown University Press.

Tomasello, M. (1998). Emulation learning and cultural learning. *Behavioral and Brain Sciences*, 21:703–704.

Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Harvard University Press.

Tonkes, B. and Wiles, J. (2002). Methodological issues in simulating the emergence of language. In Wray, A., editor, *The Transition to Language*. Oxford: Oxford University Press.

Wasow, T., Perfors, A., and Beaver, D. (2005). The puzzle of ambiguity. In Orgun, O. and Sells, P., editors, *Morphology and the Web of Grammar: Essays in Memory of Steven G. Lapointe*. CSLI Publications.

Whiten, A. (2005). The second inheritance system of chimpanzees and humans. *Nature*, 437:52–55.

Whiten, A., Custance, D., Gomez, J.-C., Teixidor, P., and Bard, K. (1996). Imitative learning of artificial fruit processing in children (*Homo sapiens* and chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, 110(1):3–14.

Whiten, A., Horner, V., and de Waal, M. (2005). Conformity to cultural norms of tool use in chimpanzees. *Nature*, 437:737–740.

Wray, A. (2000). Holistic utterances in protolanguage: The link from primates to humans. pages 285–302. Cambridge: Cambridge University Press.

Zipf, G. (1949). *Human Behavior and the Principle of Least Effort*. New York: Addison-Wesley.

Zuidema, W. (2003). How the poverty of the stimulus argument solves the poverty of the stimulus argument. In Becker, S., Thrun, S., and Obermayer, K., editors, *Advances in Neural Information Processing Systems 15 (Proceedings of NIPS'02*. MIT Press.