



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

## Complex Adaptive Systems and the Origins of Adaptive Structure: What Experiments Can Tell Us

### Citation for published version:

Cornish, H, Tamariz, M & Kirby, S 2009, 'Complex Adaptive Systems and the Origins of Adaptive Structure: What Experiments Can Tell Us: Special issue on Language as a complex adaptive system', *Language Learning*, vol. 59, no. s1, pp. 187-205. <https://doi.org/10.1111/j.1467-9922.2009.00540.x>

### Digital Object Identifier (DOI):

[10.1111/j.1467-9922.2009.00540.x](https://doi.org/10.1111/j.1467-9922.2009.00540.x)

### Link:

[Link to publication record in Edinburgh Research Explorer](#)

### Document Version:

Peer reviewed version

### Published In:

Language Learning

### Publisher Rights Statement:

This is an author's accepted manuscript of the following article: Cornish, H., Tamariz, M. & Kirby, S. 2009, "Complex Adaptive Systems and the Origins of Adaptive Structure: What Experiments Can Tell Us: Special issue on Language as a complex adaptive system", in *Language Learning*. 59, s1, p. 187-205, which has been published in final form at <http://dx.doi.org/10.1111/j.1467-9922.2009.00540.x>

### General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

### Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# Complex adaptive systems and the origins of adaptive structure: what experiments can tell us

Hannah Cornish, Monica Tamariz, Simon Kirby

Language Evolution and Computation Research Unit, The University of Edinburgh

## 1. Introduction

The position paper in this volume (Beckner et al, 2009) sets out a powerful motivating picture of language as a complex adaptive system. It presents us with a way of thinking not only about the dynamics of language as we know it now, but also the emergence of language in the first place. More specifically, in this paper, we suggest that the very fact that language persists through multiple repeated instances of usage can explain the origins of key structural properties that are universally present in language. Because of this, taking a complex adaptive systems perspective on language lifts the burden of explanation for these properties from a putative richly structured domain-specific substrate, of the sort assumed by much of generative linguistics (e.g., Chomsky, 1965). Ultimately, this alters our view of what biological evolution must provide in order to get language off the ground.

Much of the work over the past 20 years or so in modelling the evolution of language has taken this complex adaptive systems perspective (see, e.g. Kirby 2002a; Steels 2003; Brighton et al 2005 for review). One particular strand of work has focused on the adaptation of language through a repeated cycle of learning and use within and across generations, where adaptation is taken to mean a process of optimisation or *fitting* of the structure of language to the mechanisms of transmission (Kirby 1999).<sup>1</sup>

A particular subset of models have looked extensively at the impact of repeated learning on the process of emergence. They investigate how a form of cultural evolution known as *iterated learning* affects the structure of language (e.g. Batali, 1998; Kirby, 1999; Hurford, 2000; Kirby & Hurford, 2002; K. Smith, 2002; Brighton, 2002; Zuidema, 2003; A. Smith 2005; Vogt, 2005; Griffiths & Kalish, 2007; Kirby, Dowman & Griffiths, 2007). In these models, each agent (i.e., simulated individual) must acquire a set of (initially random) mappings between meanings and signals by observing the behaviour of agents in the previous generation. Once this mapping is acquired, the learner becomes a teacher, and the process repeats. Crucially there is a bottleneck in the transmission process which puts pressure on the system to be generalisable (Deacon, 1997). This bottleneck models the data-sparsity present in real language acquisition, and is typically enforced in the simulations by the learner only being exposed to signals for a sub-set of the total meanings during training.

---

<sup>1</sup> Underlying this work is a typically unstated assumption that modern languages are already optimised for transmission (i.e. all extant languages are both learnable by children and meet the expressive needs of their users), thus further change is driven not so much by inherent properties of linguistic variants, but rather sociolinguistic factors (e.g. Croft 2000). However, when looking at the origins of language, we necessarily need to consider a different state of affairs - where language has not yet reached equilibrium and the *inherent* structural properties of linguistic variants are relevant. A related point is the likelihood that intergenerational transmission is less important in ongoing language change than it is in language emergence. Where social status, for example, is the primary driving force behind selection of variants, the impact of learners' innovations is likely to be lower than where those innovations actually make language transmissible at all.

Overall, two consistent conclusions have been drawn from this computational research: that over time, iterated learning ensures languages evolve to (1) become easier to learn, and (2) become more structured. These two facts are not unrelated - one of the ways in which a language can evolve to become more learnable is by becoming structured. This is because there are only two ways to survive the transmission bottleneck: be heard (and remembered) by the next generation, or be easily inferable from what is heard. This latter solution can only occur when there is some kind of regularity to be exploited in the system. The exact form this regularity takes can vary, which is something we explore later.

The regularity that emerges gradually in the computational simulations justifies our use of the term “adaptive” in this case. This is because the kinds of linguistic structure that evolve show the hallmarks of *apparent design*. For example, in some models (e.g., Batali, 2002; Kirby, 2002b), recursive compositional syntax evolves that clearly enable the simulated agents to successfully convey meanings in an open-ended way. This kind of adaptive structure in language might lead researchers to conclude that it must reflect innate constraints that are the result of biological evolution by natural selection (e.g., Pinker & Bloom, 1990). However, this conclusion is not justified. In most of these models, there is no biological evolution. Indeed, individuals are essentially clones throughout. Rather, the adaptation arises purely from the iterated learning process itself. Language transmission is a complex adaptive system.

Recently, we developed a method for studying this process of adaptive evolution in the laboratory, extending experimental studies of iterated learning in the non-linguistic domain by Griffiths & Kalish (2007) and Kalish et al. (2007). By combining two experimental techniques: artificial language learning (e.g., Esper, 1925, 1966; Saffran et al., 1996; Gomez & Gerkin, 2000; Fitch & Hauser, 2004), and diffusion chains (e.g., Bartlett, 1932; Bangerter, 2000; Mesoudi et al. 2006; Whiten et al., 2005; Horner et al., 2006), we were able to track the evolution of a miniature language over “generations” of experimental participants from an initially random, unstructured state, to one showing clear evidence of adaptive structure (Kirby, Cornish & Smith, 2008).<sup>2</sup> In this paper, we provide a new analysis of the results of this study to examine in more detail the way structure emerges as a result of competition between linguistic variants.

## 2. Human Iterated Learning: an overview

Before we move onto the details of the studies, it is necessary to familiarise ourselves with the general methodology and key parameters of the experiments that follow. A participant is trained on an ‘alien’ language consisting of a set of meanings (usually presented as pictures) paired with signals (a string of letters, or possibly sounds) drawn from a finite set. After being trained on some proportion of these meanings, the participant is then presented with a series of meanings without signals, and asked to provide the correct description in the alien language. These meanings and signals are recorded, and become the new set of training pairs for the next participant, who forms the next ‘generation’ of the chain. This procedure is repeated, until the chain is complete (i.e. until the desired number of generations has been reached).

Participants involved in the study are only asked to learn the language as best they can: they are not told anything about the iterated nature of the study or that their responses will be given to future participants. During each training round, participants are shown a picture drawn at random from the set of meanings, and below it, a string of letters which they are told represents how the

---

<sup>2</sup> There are other experimental approaches to the origins of language, such as Galantucci (2005) and Selten & Warglien (2007), but note that these rely on participants intentionally and consciously designing a communicative system. Our interest is in whether the adaptive structure of language can arise without intentional design.

alien would describe that picture in its own language. Training occurs via a computer, and each exposure is timed to ensure no training item (meaning-signal pair) is seen for longer than any other, and continues until all training items have been seen. During the final test, the participant is shown each picture in the language once, one after another, and asked to type in the missing descriptions. These responses are then randomly sampled from to generate the new training items for the next generation.

Clearly, this experimental setup represents a highly simplified idealisation of the real process of linguistic transmission. In particular, the population model is the simplest that we could construct (in line with the other diffusion chain experiments mentioned previously). Three parameters characterise possible population models: direction of transmission (vertical or horizontal), the size of the population, and who learns from whom (network structure). For the rest of this paper we focus on just one scenario - vertical transmission, involving ten people, with each person learning from just one other person. However, it is important to remember that there are many other scenarios that could be explored within this framework.

### 3. Learnability, expressivity and adaptation

As stated in the introduction, the main finding to have emerged over the past decade or so of research into this area is that languages themselves adapt to be better learnable and transmissible by us over time (see e.g., Christiansen & Chater, 2008, for a review). However, it should be recognised that this pressure toward greater learnability must be tempered somewhat in order for structure to emerge. The reason for this is simple: the most easily learnable language might be one in which there is one word for everything (or possibly, no words at all). It is only when we also have a pressure for expressivity, for meanings to actually be distinguished from one another, that we are likely to see the emergence of structure.

The first application of this new experimental methodology set about investigating this tension between expressivity and learnability (Kirby, Cornish & Smith, 2008). In this study, the meaning-space consisted of 27 pictures showing a scene which varied along three features and three values: colour of object (blue, black, red), shape of object (circle, triangle, square), and a dotted line indicating the movement of object (bouncing, spiralling, moving horizontally). Two different experimental conditions were explored, with four chains of ten people in each. In one condition there was a "hidden" pressure for each of the meanings in the meaning-space to be expressed uniquely - participants' input was filtered in such a way as to ensure they never perceived different meanings with the same signal. In the other there was no such pressure. Participants could not be aware of which experimental condition they were in.

The chains in each condition both began with random initial languages, and a transmission bottleneck was imposed by exposing each generation with just half (14) of the meaning-signal pairs during training (the particular meanings that they would be exposed to were chosen randomly each generation). Example 1 shows a sample of the initial randomly generated language in one of the chains to illustrate what is meant by the claim that they are unstructured with respect to their meanings.<sup>3</sup> In spite of the fact that these meanings in the world are similar (triangles of every colour that either move horizontally or in a spiral), the signals used to describe them are all idiosyncratic, with no consistently repeating sub-parts.

(1)    a)    kapihu                                  b)    luki

---

<sup>3</sup> The glosses here are given as English words - recall that in the experiment, visual stimuli were used. This example is taken from Chain 3 in Experiment 2 in Kirby, Cornish & Smith (2008).

	‘black triangle horizontal’		‘black triangle spiral’
c)	humo	d)	namola
	‘blue triangle horizontal’		‘blue triangle spiral’
e)	lahupiki	f)	lumoka
	‘red triangle horizontal’		‘red triangle spiral’

After training, participants were tested on all 27 meanings, and it is from this output set that the new training set is sampled for the participant in the next generation.

The main findings can be summarised as follows (see Kirby, Cornish & Smith (2008), for more details). Firstly, by looking at the learning errors made between adjacent generations, it was shown that the languages in both conditions were being acquired significantly more faithfully towards the end of the chains than they were at the beginning. Secondly, this increase in learnability over time occurred as a result of the languages becoming more structured over time.

What is interesting about this last fact however, is that the way in which the languages were structured differed markedly between the two experimental conditions. In the first condition, where there was no filtering of the participants’ input, systems emerged that were characterised by *underspecification*. This involved a reduction in the total number of distinct signals, introducing ambiguity with respect to the meanings. However, this ambiguity was not complete, as it did not affect all meaning dimensions. In one chain for instance, a system emerged (of which a sample is reproduced as Example 2) whereby everything that moved horizontally was called **tuge**, everything that moved in a spiral was named **poi**, and there was a three way distinction of bouncing items dependent on shape: for bouncing squares, **tupim**, bouncing triangles, **tupin** and bouncing circles, **miniku**. This system proved to be highly adaptive in the sense that, once it emerged, it was stable and faithfully acquired by subsequent generations without error<sup>4</sup>.

(2)	a)	tuge	b)	poi
		‘black triangle horizontal’		‘black triangle spiral’
	c)	tuge	d)	poi
		‘blue triangle horizontal’		‘blue triangle spiral’
	e)	tuge	f)	poi
		‘red triangle horizontal’		‘red triangle spiral’

As Kirby, Cornish & Smith (2008) point out, underspecification is not an unusual feature of human languages, but taken to extremes it would lead to an inexpressive and communicatively dis-

<sup>4</sup> This is not a trivial result considering the rather narrow bottleneck applied during training meant that each generation was being trained on a (different) random subset of half of the total language.

functional language (albeit one that would be easy to learn). The second experimental condition, whereby items were removed from a participant's input if they should lead to the same string being assigned to more than one meaning, was designed to introduce a hidden pressure against underspecification. With this modification in place, the systems that emerged appear much closer to what we might expect a communicatively useful system to look like. These systems were characterised by *compositionality*, whereby the meaning of a given string could be inferred by the meaning of sub-parts of that string (morphemes) and the way they are put together. Example 3 again shows a sample of this<sup>5</sup>.

(3) a)	nekeki	b)	nekipilu
	'black triangle horizontal'		'black triangle spiral'
c)	lakeki	d)	lakipilu
	'blue triangle horizontal'		'blue triangle spiral'
e)	raheki	f)	rahopilu
	'red triangle horizontal'		'red triangle spiral'

These results are very exciting, as they experimentally verify the main findings to have emerged from computational models of iterated learning for the first time: that languages adapt purely by virtue of transmission through iterated learning. Moreover, the kind of adaptation is determined in part by constraints placed on the transmission of the languages about which participants could not be aware. However, whilst it has been shown that the languages in these experiments *do* adapt, it has not yet been established *how* they adapt. It is to this question that we now turn.

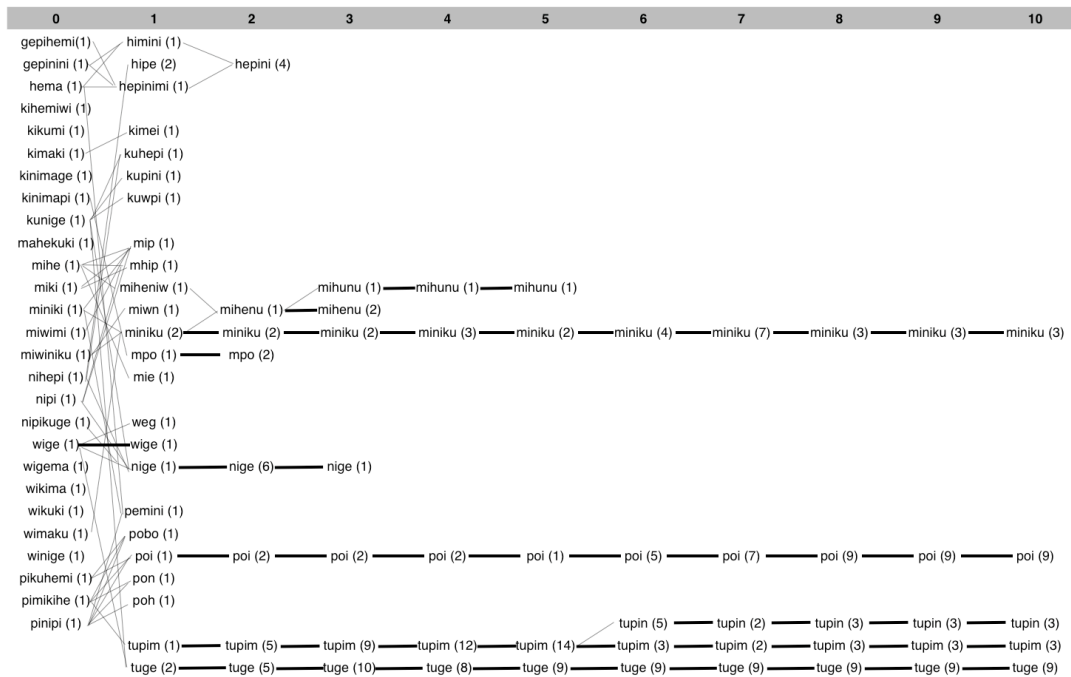
### The evolution of signals during iterated learning

In this section we will focus on the utterances, leaving aside the meanings for the moment, and construct phylogenies demonstrating the evolution of linguistic forms over iterations. We used one of the languages (part of which was reproduced above in example 2), taken from Kirby, Cornish & Smith (2008) to construct the coalescent tree shown in Figure 1. These trees are a standard way to represent phylogenetic descent in evolutionary biology (Hein et al., 2005; Barton, 2007), although here we have amended them to also include frequency information in brackets. Bold lines show perfect replication of an utterance, whilst other lines show possible relationships of descent between utterances across generations.

As we can see in Figure 1, the number of different utterances decreases over time as we start to observe perfect replication of select utterances, along with a general tendency for utterances to become shorter. In the early history of this language, the process of transmission is principally one of generating new recombinations of signal sub-strings. We only observe one instance of replication

<sup>5</sup> Taken from generation 9, chain 3, experiment 2 in Kirby, Cornish & Smith (2008). Note that while colour and motion are consistently expressed (ne for black, la for blue, ra for red, ki for horizontal and pilu for spiral), shape is more haphazardly encoded (ke when blue/black and horizontal, ki when blue/black and spiral, he when red and horizontal, and ho when red and spiral).

of a whole utterance, but many replications of parts of the utterances, such as unigrams or bi-grams, and even larger n-grams. For example, the introduction of the form *miniku* in generation 2 could be the result of a blend between *miniki* and *miweniku*<sup>6</sup>. There is still a lot of variation in the language at this point. In the final generations however, the frequencies of the few remaining units stabilise around multiples of three, suggesting adaptation to a meaning space containing three dimensions.



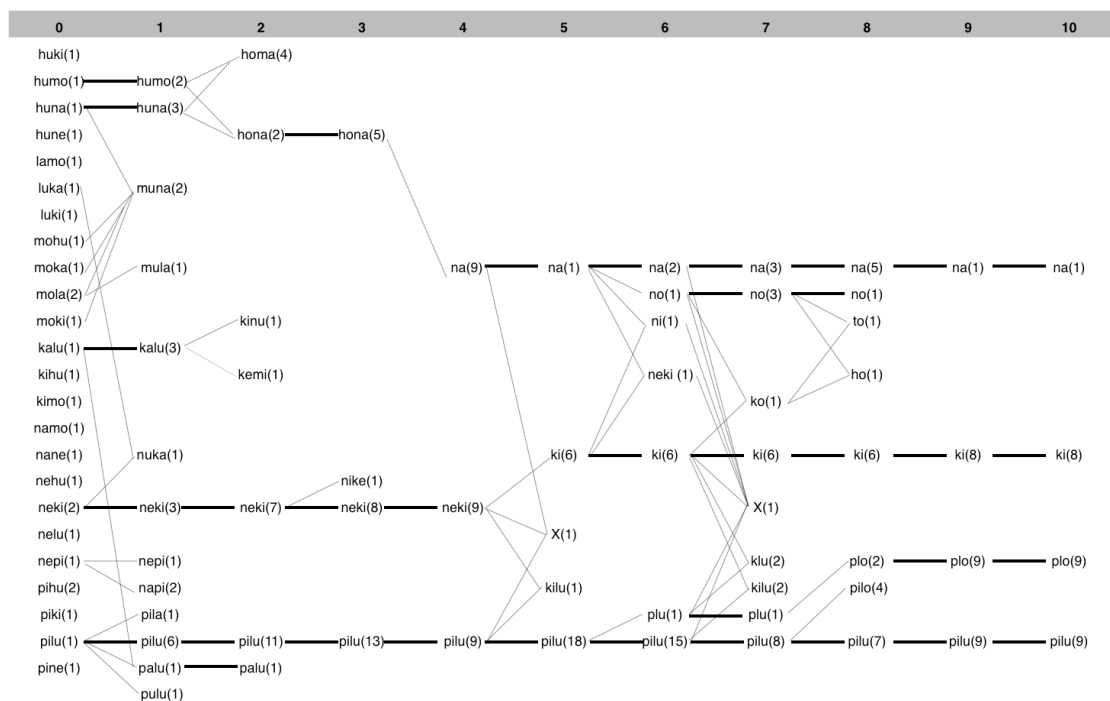
**Figure 1.** Coalescent tree showing lineages of signals for all 27 items over generations of one of the languages obtained by Kirby, Cornish & Smith (2008) exhibiting systematic underspecification. Columns correspond to generations; horizontal bold lines indicate perfect replication of the whole signal; all other lines indicate some of the possible relationships of descent between signals that share some features. Numbers shown in brackets indicate the frequency with which variants were produced at each generation. The number of variant types decreases over time, although the number of tokens remains fixed at 27 throughout. Amongst these surviving variants there are clear relationships of descent, sometimes with modification. The frequency information is suggestive of the fact that signal variants may be adapting to express a meaning space composed of multiples of three.

In the case of the language in Figure 1, given the non-decomposable utterances that survived into the final stable system, it was appropriate to analyse replication at the level of the whole utterance.

<sup>6</sup> It is perhaps interesting to note that the investigation of this type of phenomenon, historically referred to as analogical change, was what prompted the very first application of this methodology by Esper in 1925.

However, in a compositional system, the meaning of a complex utterance is a function of the meanings of the elements of the utterance and the way they are arranged. The tree in Figure 1 illustrates adaptation of the whole signals to the structure of the meaning space; in a compositional language, we expect the same phenomena to occur but this time at the level of signal elements. We will now quantify compositionality using a different language (part of which is shown in example 3 above) from Kirby, Cornish & Smith (2008).

First we need to segment the signals into element units. To do this we first examined the language of the final participant in the chain to find the most parsimonious segmentation of the strings into elements that corresponded to aspects of the meanings (for example, “the signal endings reliably encode motion”, or “signal-initial ‘la’ consistently encodes colour blue”). This resulted in each string being divided into three sub-strings, and this segmentation pattern was carried back to all previous generations in order to allow for a consistent analysis. Figure 2 shows the coalescent tree for the word-final signal element (although similar trees can be constructed for both initial and middle positions also).



**Figure 2.** Coalescent tree showing lineages of signal ending variants for all 27 items over ten generations of one of the languages obtained by Kirby, Cornish & Smith (2008), referred to in Example 3. Numbers shown in brackets indicate the frequency with which variants were produced at each generation. The number of variant types decreases over time, although the number of tokens remains fixed at 27 throughout. Amongst these surviving variants there are clear relationships of descent, sometimes with modification. The frequency information is suggestive of the fact that signal variants may be adapting to express a meaning space consisting of three meaning elements (see generations 4, 9 and 10).



As before, we observe a marked reduction in the number of variants over time, as just a few become selected to be re-used more often. Furthermore, we can see that the variants that appear at each generation are not random; we can trace the genealogy of the surviving variants back in time. Even over this minute time-scale, many of the changes observed appear to follow paths that are well attested in cases of natural language change, such as single segment replacements (nepi → napi; pilu → pilo), reductions (hona → na, neki → ki, pilu → plu), metathesis (neki → nike), and blends (humo & huna → homa & hona; na & ki → neki).

It is significant to notice that at generation four we have three variants (na, neki, pilu) each with a frequency of nine, and that for the final two generations this pattern repeats (now for variants ki, plo, pilu) broken only by a single instance of na. This again suggests that these lineages are adapting to a three-element meaning space. Obviously we know that this is indeed the case - the interesting thing is that the signals alone suggest it. In the next section, we show how we can precisely quantify regularities in the mappings between signal and meaning elements in order to objectively confirm this.

## Quantifying the emergence of compositionality

We now have an analysis of all the languages in terms of: *signal segments* - in this case, the word beginning, middle or end; *signal segment variants*, actual tokens residing in a segment position - such as pilu or ki. Similarly, we can define: *meaning elements*, aspects of meaning - motion, shape and colour; and *meaning element variants*, actual instances of a meaning element - for instance, 'blue' or 'circle' or 'bounce'.

Kirby, Cornish & Smith (2008) quantify the emergence of structure using pairwise distance correlation (Shillcock et al., 2001). This measures the extent to which similar meanings are expressed using similar forms – or more precisely, whether there is a correlation between the structure of the meaning and signal spaces. Although this is valuable in showing that structure emerges, it does not allow us to track the evolution of the compositional structure of the languages directly: as a measurement, the pairwise distance correlation is very general and cannot distinguish between compositionality and other kinds of structure (such as underspecification). Here we apply a new method of analysis to one of the chains<sup>7</sup> reported in Kirby, Cornish & Smith (2008) to tackle this problem. We use *RegMap* (Tamariz & Smith, 2008), an information-theoretical metric that combines the conditional entropy of meanings given signals and of signals given meanings and normalises the result to make it comparable across systems of different sizes. Informally, what *RegMap* (short for regularity of the mappings) does is return the degree of confidence that a signal element consistently predicts a meaning element - for instance, the degree to which we can be sure that the beginning of the signal encodes colour.

More formally,  $H(X|Y)$ , the conditional entropy, is the Shannon entropy (Shannon, 1948) but replacing  $p(x)$  with  $p(x|y)$ . The *RegMap* for a meaning element (M) and a signal segment (S) is given in Eqn. 1.

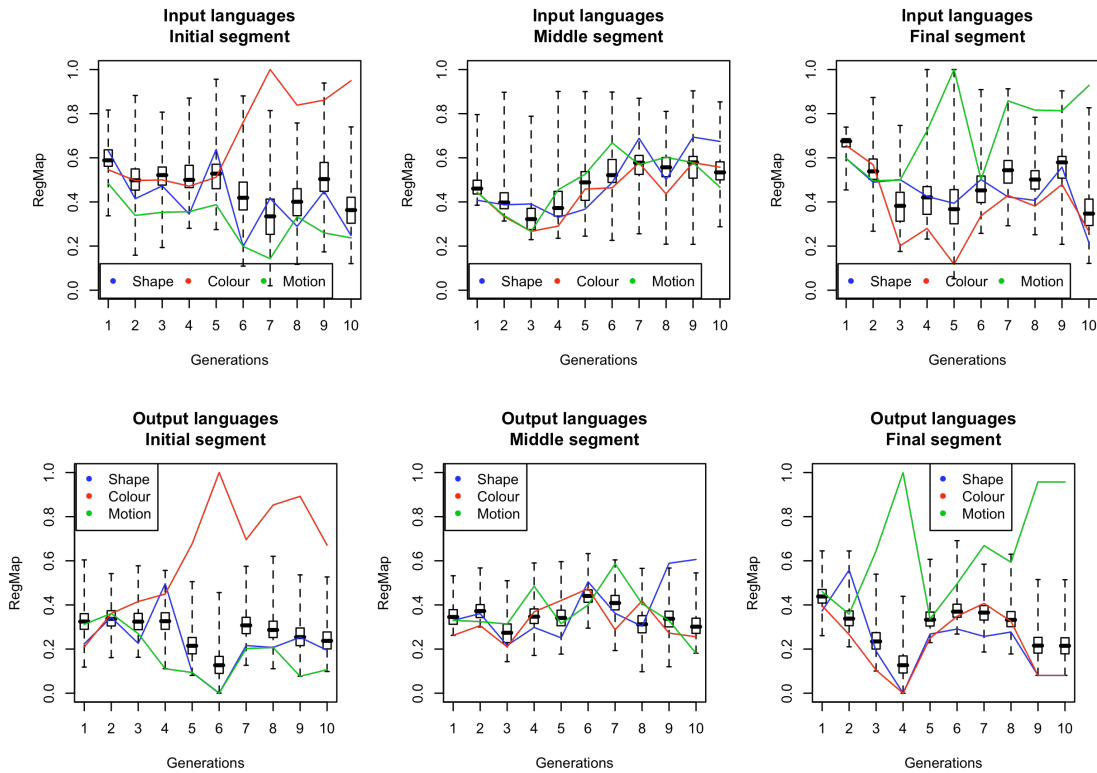
---

<sup>7</sup> Specifically, we examine chain 3 in experiment 2, but similar results can be obtained wherever compositionality clearly emerges.

$$(1) \quad \text{RegMap} = \sqrt{\left(1 - \frac{H(S|M)}{\log(n_s)}\right) \times \left(1 - \frac{H(M|S)}{\log(n_m)}\right)}$$

$H(S|M)$  is the conditional entropy of the signal segment given the meaning feature, or the uncertainty about the meaning when we know the segment. This relates to comprehension. For example, for shape and first signal segment,  $H(S|M)$  quantifies how uncertain we are on average about what shape an object is if we hear the first segment of its corresponding signal.  $H(M|S)$  is the conditional entropy of the meaning feature given the signal segment, or the uncertainty about the segment when we know the meaning. This relates to production. Still, in the case of shape and first signal segment,  $H(M|S)$  quantifies how uncertain we are on average about what first segment to produce if we know the shape of an object. The logs of  $n_m$  and  $n_s$  normalise the values between 0 and 1;  $n_m$  is the number of different meaning values (e.g. triangle, circle, square for shape);  $n_s$  is the number of different segment variants in the relevant segment position. Subtracting the conditional entropies from 1 returns levels of confidence instead of uncertainty.

Figure 3 shows the *RegMap* values for all combinations of signal and meaning elements both with and without a bottleneck for the 10 generations. The ‘input’ data shown in Figure 3 (upper) reflects the extent to which signals predict meanings in the sub-set of the language (taken from the previous generation) that was actually transmitted to the current generation, after the bottleneck was applied. The ‘output’ data shown in Figure 3 (lower) is obtained from the complete languages that participants actually produced at a given generation, before the bottleneck was applied. The significance of the obtained *RegMaps* was established with a Monte-Carlo analysis involving 10,000 randomisations of the correspondences between meanings and signals, and are shown as boxplots.



**Figure 3:** Regularity of the associations between signal and meaning elements, measured as *RegMap*, changes over time in the direction of maximising compositionality, whereby signal elements are consistently associated with distinct meaning elements. The continuous coloured lines represent *RegMap* values obtained with all nine segment-meaning feature pairs in the ten generations of a language family from Kirby, Cornish and Smith 2008, referred to in Example 3. The boxplots show the distributions of values obtained with 10,000 randomised languages. The upper graphs show *RegMap* values from the sub-set of language (taken from the previous generation) that was actually transmitted to the current generation, after the ‘bottleneck’ was applied. The lower graphs show *RegMap* values obtained from the complete languages that participants actually produced at a given generation, before the bottleneck was applied.

Focusing first on the bottom graphs, obtained from the participants’ output languages, we see that, starting off from values indistinguishable from random at generation 1, *RegMap* becomes massively increased to highly statistically significant levels; specifically, by the third generation, motion is consistently encoded by the final signal segment, by the fourth generation colour is encoded by the initial segment and by the ninth generation, shape is encoded by the middle segment (all  $p < 0.001$ ).

Second, a comparison of the input (upper) and output (lower) results in Figure 3, reveals the effect of the bottleneck. The *RegMap* values are, in the majority of cases, amplified by the bottleneck (the absolute value of *RegMap* increases). Moreover, the lower the input *RegMap*, the more likely it is to be amplified by the bottleneck. How is this happening? The answer is potentially counterintuitive; randomly occurring patterns are more likely to be perceived the smaller the system is. At least in the early generations, a sub-set drawn from a language is more likely to accidentally contain more regular patterns than the entire language. Implicit in this, and by the same token, a given sub-set will also tend to contain less counter-evidence against such patterns. This explains why we observe such a dramatic difference between the ranges shown in the boxplots in the upper and lower graphs in Figure 3. The large range of *RegMap* values in the input languages directly reflects the fact that participants are sensitive to this reduced number of observations when they are inferring the mappings between meanings and signals. Together, this accounts for the structure-generating effect of the bottleneck on language: the input to each generation is only a fraction of (and therefore tends to be more systematic than) the total output of the previous generation.

Third, the graphs show cases of competition between meanings all vying to be expressed by the same signal element. For example, motion and shape are both equally encoded in the final signal segment in the input to generation 3, but generation 3 resolves this conflict by ignoring one of the associations (shape) and amplifying the other (motion) to a significance level of  $p < 0.01$ . Conversely, we also see cases of competition between signals vying to express the same meaning: in the input of generation 5 colour is equally encoded in the initial and middle signal segments (similar absolute values and levels of significance); in this case the conflict is resolved by massively amplifying the association with the initial segment to a significance level of  $p < 0.001$ , and reducing the association with the middle one. These processes are adequately explained by the standard evolutionary mechanisms of variation, replication, and selection applied to the mappings between signals and meanings elements. Selection, in this case, can be hypothesized to be guided by perceptual and attentional biases such as higher salience of certain signal and meaning elements over others. A detailed discussion of these biases unfortunately lies outwith the scope of the present paper.

## Summary

Kirby, Cornish & Smith (2008) found that the languages that emerge through a repeated cycle of learning and production in a laboratory setting show evidence of adaptation to the bottleneck placed on their transmission. Making even minor changes to the way in which language is culturally transmitted can produce radically different types of structure. Given only a bottleneck on transmission preventing a proportion of the language from being seen by the next generation, language can adapt in such a way that ensures it is stably transmitted to future generations regardless. This however occurs at the expense of being able to uniquely refer to every meaning. When they introduced the additional pressure of having to use a unique signal for each meaning, the language once again adapted to cope with these new transmission constraints, this time by becoming compositional. Having a compositional system ensures that both signals and meanings survive the bottleneck.

Because the participants could not know which condition they were in, it is impossible that the resulting languages were intentionally designed as adaptive solutions to the transmission bottleneck. Rather, the best explanation for the result is that in these experiments, just as in the computational models, linguistic adaptation is an inevitable consequence of the transmission of linguistic variants under particular constraints on replication. The result is apparent design, but without an intentional designer.

Whereas Kirby, Cornish & Smith (2008) analysed their results at the level of whole signals and whole meanings, in this section we have developed new techniques to analyse the same results in terms of the component parts of linguistic signals. An analysis of how signal variants and their frequencies change over time showed relationships of descent with modification among them. It also suggested that signal variants are adapting to the structure of the meaning space. This intuition was verified by the application of *RegMap*, a tool designed to objectively measure compositionality. Using this method, we showed that individual signal elements come to encode individual meaning elements, while the whole system evolves to avoid ambiguity (i.e. more than one meaning being encoded in the same signal element or vice versa). Moreover, we were able to more precisely describe the role of the bottleneck in bringing about compositionality: the smaller sub-sets sampled as inputs to the next generation may locally contain more systematicity than the entire language. Iterating this learning process using these smaller samples therefore provides a platform that allows systematic patterns to be noticed, remembered and replicated preferentially, thereby allowing them to gradually accumulate in the language as a whole.

It seems clear from all of this that firstly, cultural transmission alone is capable of explaining the emergence of languages which exhibit that appearance of design, and secondly, experimental studies of the iterated learning of artificial languages are a potentially useful methodological tool for those interested in studying cultural evolution.

## 4. Conclusion

This paper has extended previous work on iterated language learning experiments by showing, using data obtained from an earlier study, exactly how compositional structure emerges over time as a result of cultural transmission. Using a recently developed analytical technique, which calculates the regularity of mapping between signal and meaning elements (Tamariz & Smith, 2008), we were able to precisely quantify changes in the language's ability to systematically encode such associations between meaning and signal components. From this we were able to explain the amplification effect the bottleneck seems to have on systematicity in language, arguing that the sampling of smaller sub-sets of the language for training input to the next generation tends to make weaker patterns that are not visible at the level of the entire language appear stronger locally.

One obvious criticism of the experimental work described here is that it necessarily involves participants who already speak a language. As such, can it tell us anything about the original evolution of language as we are claiming? The sceptical position might be that we are simply seeing the evolution of structure that reflects the native language of the participants as opposed to any adaptive logic of the iterated learning process itself. This criticism faces a number of problems, however. Most importantly, the experimental results are backed up by the computational simulations and mathematical models surveyed in the introduction. In these models we can be sure there is no influence of prior language, since the models have none initially. Furthermore, the structure that arises depends on aspects of the transmission bottleneck that are hidden from our participants (given our two experimental conditions) and the particular properties of the language appear more dramatically shaped by these than any similarity to the language of the participants. The most parsimonious explanation, then, is that we are seeing adaptation to the transmission bottleneck rather than an emerging simple first-language influence. However, a more subtle point can be made: we fully expect that language evolution through iterated learning will involve adaptation to all aspects of the transmission bottleneck, and this will include the biases of language learners. In our experiment, participants bring to bear a mixture of biologically basic biases and those that arise from their acquired cultural heritage. We can see no principled way to separate these out. This means that our experiments should not be taken to be a “discovery procedure” for uncovering our evolutionary ancient learning biases but rather as a tool for understanding the fundamental adaptive dynamics of the cultural transmission of language by iterated learning.

We started this paper by noting that a complex adaptive systems perspective shifts the burden of explanation away from a richly structured domain-specific innate substrate for language in our species. Although we have talked a lot about linguistic structure as an adaptation, this is adaptation by the language itself rather than biological evolution of the faculty of language. The relevant explanatory mechanisms relate to *cultural* as opposed to *natural* selection. But of course, this does not mean that biology is irrelevant to the evolution of language.

Rather than seeking evolutionary explanations for innate constraints that determine language structure, the work presented in this paper strongly suggests a different approach. The iterated learning models we base our experiments start with agents who can a) learn complex signals, and b) infer complex meanings. Humans belong to an unusual set of species, called the “vocal learners” (Jarvis 2004) that can learn sequential signals (others include most notably the songbirds). We are also unusually adept in inferring intentionality (Tomasello et al. 2005). By taking into account the power of language as a complex adaptive system to generate structure itself, future work on the biological evolution of language in our species should focus on how we came to have these two crucial pre-adaptations for language. Without the combination of vocal learning and meaning inference, iterated learning of the sort we are studying would not be possible at all (Okanoya 2002). Once they are in place, on the other hand, the emergence of structure is inevitable.

## Acknowledgements

Hannah Cornish is funded by an ESRC postgraduate studentship. Monica Tamariz was funded by a Leverhulme Early Career Fellowship. Monica Tamariz and Simon Kirby are supported by ESRC Grant Number RES-062-23-1537 and by AHRC Grant Number AH/F017677/1. The work presented here benefited from discussion among the members of the University of Edinburgh’s Language Evolution and Computation Research Unit, and constructive comments from two anonymous referees.

## References

- 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.
- Barton, N. H. (2007). *Evolution*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Batali, J. (1998). Computational simulations of the emergence of grammar. In Hurford, J., Studdert-Kennedy, M., and Knight, C., editors, *Approaches to the Evolution of Language: Social and Cognitive Bases*. Cambridge University Press.
- Batali, J. (2002). The negotiation and acquisition of recursive grammars as a result of competition among exemplars. In Briscoe, T., editor, *Linguistic Evolution through Language Acquisition: Formal and Computational Models*, chapter 5. Cambridge University Press.
- Beckner, C., Blythe, R., Bybee, J., Christiansen, M.H., Croft, W., Ellis, N., Holland, J., Ke, J., Larsen-Freeman, D., & Schoenemann, T. (2009). Language is a complex adaptive system. *Language Learning*.
- Brighton, H. (2002). Compositional syntax from cultural transmission. *Artificial Life*, 8(1):25–54.
- Brighton, H., Smith, K., and Kirby, S. (2005). Language as evolutionary system. *Physics of Life Reviews*, 2:177-226.
- Bybee, J. L. (2006). *Frequency of use and the organization of language*. Oxford University Press.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. MIT Press.
- Christiansen M.H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Science*, 31(5):489–509.
- Croft, W. (2000). *Explaining Language Change: an evolutionary approach*. Longman.
- Deacon, T. W. (1997). *The Symbolic Species: The Co-evolution of Language and the Brain*. W.W. Norton.
- Esper, E. A. (1925). *A technique for the experimental investigation of associative interference in artificial linguistic material*. Philadelphia: Linguistic Society of America.
- Esper, E. A. (1966). Social transmission of an artificial language. *Language*, 42(3): 575–580.
- Fitch, T., & Hauser, M. (2004). Computational Constraints on Syntactic Processing in a Non-human Primate. *Science*, 303: 377-380.
- Francis, N. & Kucera, H. (1982). *Frequency Analysis of English Usage: Lexicon and Grammar*. Hoghton Mifflin.
- Griffiths, T. L., & Kalish, M. L. (2007). Language evolution by iterated learning with Bayesian agents. *Cognitive Science*, 31(3):441–480.
- Galantucci, B. (2005). An experimental study of the emergence of human communication systems. *Cognitive Science*, 29:737–767.
- Gómez, R. L., & Gerken, L. A. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences*, 4(5): 178-186.

- Hein, J., Schierup, M. H., & Wiuf, C. (2005). *Gene genealogies, variation and evolution: a primer in coalescent theory*. Oxford University Press, Oxford.
- Horner, V., Whiten, A., Flynn, E., & de Waal, F. B. M. (2006). Faithful replication of foraging techniques along cultural transmission chains by chimpanzees and children. *PNAS*, 103(37):13878–13883.
- Hurford, J. (2000). Social transmission favours linguistic generalization. In Chris Knight, J. R. H. and Studdert-Kennedy, M., editors, *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*, pp. 324–352. Cambridge University Press, Cambridge.
- Jarvis, E. (2004). Learned birdsong and the neurobiology of human language. *Annals of the New York Academy of Sciences*, 1016:749–777.
- Kalish, M. L., Griffiths, T.L., & Lewandowsky, S. (2007) Iterated learning: Intergenerational knowledge transmission reveals inductive biases. *Psychonomic Bulletin and Review*, 14(2):288–294.
- Kirby, S. (1999). *Function, Selection and Innateness: the Emergence of Language Universals*. Oxford University Press, Oxford.
- Kirby, S. (2001). Spontaneous evolution of linguistic structure: an iterated learning model of the emergence of regularity and irregularity. *IEEE Transactions on Evolutionary Computation*, 5(2):102–110.
- Kirby, S. (2002a). Natural language from artificial life. *Artificial Life*, 8(2):185–215.
- Kirby, S. (2002b). Learning, bottlenecks and the evolution of recursive syntax. In Briscoe, T., editor, *Linguistic Evolution through Language Acquisition: Formal and Computational Models*, chapter 6, pages 173–204. Cambridge University Press.
- Kirby, S., & 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*, chapter 6, pages 121–148. Springer Verlag, London.
- Kirby, S., Dowman, M., & Griffiths, T. (2007) Innateness and culture in the evolution of language. *Proceedings of the National Academy of Sciences*. 104:5241–5245.
- Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: an experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*. 105:10681–10686.
- Mesoudi, A., Whiten, A., & Dunbar, R. (2006). A bias for social information in human cultural transmission. *British Journal of Psychology*, 97:405–423.
- Okanoya, K. (2002). Sexual display as a syntactical vehicle: the evolution of syntax in birdsong and human language through sexual selection. In Wray, A., editor, *The Transition to Language*, pages 46–63. Oxford University Press.
- Pinker, S. and Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences*, 13(4):707–784.
- Saffran, J., Aslin, R., & Newport, E. (1996). Statistical learning by 8-month olds. *Science*, 274:1926–1928.
- Selten, R., & Warglien, M. (2007). The emergence of simple languages in an experimental coordination game. *PNAS*, 104(18):7361–7366.

- Shannon, C. E. (1948). A mathematical theory of communication. *Bell Systems Technical Journal*, 27:379-423 & 623-656.
- Shillcock, R., Kirby, S., McDonald, S., & Brew, C. (2001). Filled pauses and their status in the mental lexicon. In *Proceedings of the 2001 conference on Disfluency in Spontaneous Speech*.
- Smith, A. (2005). The inferential transmission of language. *Adaptive Behavior*, 13(4):311–324.
- Smith, K. (2002). The cultural evolution of communication in a population of neural networks. *Connection Science*, 14(1):65–84.
- Steels, L. (2003). Evolving grounded communication for robots. *Trends in Cognitive Science*. 7(7), July 2003, pp. 308-312.
- Tamariz, M., & Smith, A. D. M. (2008). Regularity in Mappings Between Signals and Meanings. In Smith, A. D. M., Smith, K. and Ferrer i Cancho, R., editors. *The Evolution of Language: Proceedings of the 7th International Conference (EVOLANG7)*, pages 315–322. World Scientific.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28, 675 - 691.
- Whiten, A., Horner, V., & de Waal, M. (2005). Conformity to cultural norms of tool use in chimpanzees. *Nature*, 437:737–740.
- Vogt, P. (2005). The emergence of compositional structures in perceptually grounded language games. *Artificial Intelligence*, 167:206–242.
- 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.