

The Evolution of Linguistic Replicators

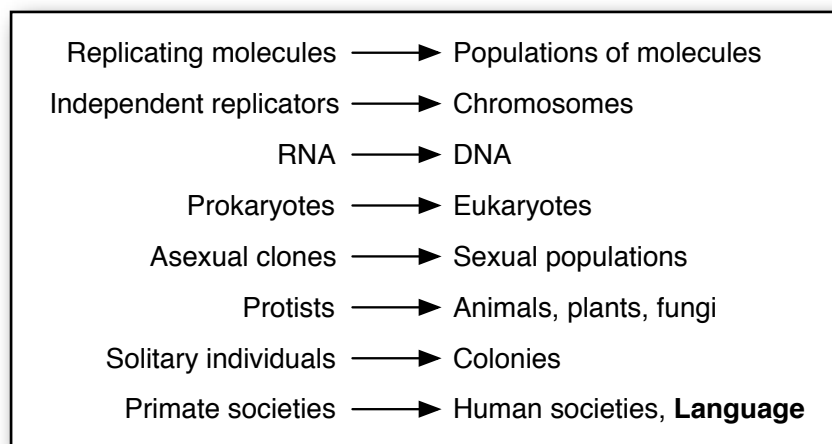
Simon Kirby

School of Philosophy, Psychology & Language Sciences
University of Edinburgh

www.ling.ed.ac.uk/~simon

1. Introduction: the Major Transitions

In their thought-provoking survey, Maynard Smith & Szathmáry (1995) [MS&S] propose a series of “major transitions” in the evolutionary history of life. These take us from the emergence of populations of replicating molecules out of solitary replicating molecules as the earliest transition, all the way to the most recent, the emergence of human language-using societies out of primate a-lingual ones. In the introduction to their work, they list eight transitions in particular:



This list is not merely a rag-bag of interesting evolutionary developments, of course. The insight that MS&S had was that there were recurring themes and commonalities among these crucial points in evolutionary history. Despite their apparent diversity, by considering these transitions together we can reveal underlying similarities such that advances in understanding any one can deepen our understanding of the others. In this chapter, I will suggest that we can extend the same logic to the evolution of language and find similar transitions which are remarkably similar in kind to the ones that MS&S discuss.

But first, I will quickly survey some of the common themes that recur throughout the major transitions:

1. **Loss of independent replication:** Examples given by MS&S include the origins of social groups such as in the social insect species in which individuals can only survive as part of a group. Another example is the origin of the eukaryotes. Ancestors of mitochondria and chloroplasts, which now replicate only as part of a cell, were once independently viable.
2. **Division of labour:** An example of this feature of transitions is given by MS&S characterisation of a transition from a hypothetical “RNA world” where RNA worked as both carrier

of genetic information and as a catalyst, to the DNA world in which DNA carries the code, but other functions are carried out by proteins.

- 3. New ways of transmitting information:** There are many examples of this in MS&S's book and in some sense it is what underpins the whole idea of evolutionary transitions. With the origins of RNA, the origins of the genetic code, and so on we see the emergence of important new systems of inheritance.

It is striking that MS&S put language on their list as part of the last transition they discuss. Why have they done this? The answer is that they feel language is a significantly new system of information transmission. Language, for MS&S, is part of the last transition because it supports human cultural transmission. Cultural information that humans possess can be *replicated* in other humans through the medium of language. Szathmáry (2000) refers to the units of information passed-on in this way as human memes, and argues that these are a new type of replicator that could only be possible with language.

Memes are an example of a *phenotypic replicator* in Szathmáry's (2000) terminology. They are transmitted without direct copying of a code. Other examples of phenotypic replication in nature include *prions*, which are proteins with have an infective phenotype that replicates by re-shaping proteins it comes into contact with directly – again, without use of a separate copied code. What makes *human* memes supported by language special for MS&S, and qualitatively different from what has gone before, is that they have *unlimited heredity*:

“A crucial distinction is between systems of ‘limited heredity’, in which only a few distinct states can be transmitted, and systems of ‘unlimited heredity’, capable of transmitting an indefinitely large number of messages.” (Maynard Smith & Szathmáry, 1995:13)

So, in summary, the emergence of language saw the appearance of a qualitatively new kind of transmission of cultural information in the primate lineage. Prior to this, cultural information was transmitted via phenotypic replicators with limited heredity, afterwards we made the transition to unlimited phenotypic replication.

2. Linguistic replication

It is perhaps striking that Maynard Smith & Szathmáry's (1995) work has not made a more significant impact on the evolutionary linguistics community, let alone linguistics itself. Part of the problem may be that it is unclear what placing language among the major evolutionary transitions actually tells us about the evolution of language. One approach might be to pursue an adaptationist account whereby our capacity for language was selected for on the basis of its meme-carrying capacity. But it is unclear how such an account would be spelled out, beyond the simplistic but common notion that language must have evolved “for communication” (see, e.g., discussion in Pinker & Bloom (1990) and its associated commentaries). The most obvious difficulty is in determining how a capacity for sharing information not only among others in a group, but also to future generations could be promoted by selection operating on individuals.

Here I want to propose another way of looking at the relationship between the evolution of language and the major transitions. If you consider an utterance, what information does that utterance convey? Most obviously, it conveys *semantic* information, and ultimately this is the information that MS&S are talking about in their discussion of unlimited cultural heredity. In other words, utterances carry *content*. But there is another kind of information that can be conveyed by any linguistic production, and that is information about the linguistic system itself. Along with content comes *form*. When I produce the sentence “these berries are good” I may be

propagating cultural information about the edibility of items in the environment via the content of the sentence. At the same time I may also be propagating information about the construction of sentences in my language: its phonology, its lexical entries, its syntactic structure and so on.¹ Obviously, the relevance of these two kinds of information transmission will vary depending on the listener. Children in the process of language acquisition will be making the most significant use of the latter kind, although the process of language change is driven to a large part by ongoing adaptation of adults' linguistic representations throughout life (see, e.g., Croft 2000, for discussion).

So, language is not only a mechanism for cultural replication, it is *itself* culturally² transmitted. By transmitting information about their own construction, utterances are part of a system of linguistic replication. There has been a great deal of debate in the literature about the best way to characterise language change in broadly Darwinian terms, and a thorough review of this debate is beyond the scope of this chapter (see, e.g., Croft 2000; Ritt 2004; Kirby 1999; McMahon 1994). Suffice it to say that there are a number of different ways of applying ideas such as replication and selection to the linguistic system, each with their own plausibility and merits. Here, I want to relate linguistic transmission to Szathmáry's (2000) characterisation of replicators.

One way of thinking about the process of language acquisition is in terms of *reverse engineering*. This is a process whereby an engineering team will attempt to copy a competitor's product by inferring the (hidden) inner workings of that product through careful observation of its external behaviour. The goal of reverse engineering is not so much to reproduce just the observed behaviours, but to generalise these to all possible behaviours of the product being copied. This is likely to result in a copy of the inner working of the product, but this is not guaranteed. Reverse engineering is appropriate when the underlying mechanism is hidden and only external behaviour is accessible. This is exactly the case for language acquisition: children cannot directly observe the state of the (neural) mechanisms that generate adult linguistic behaviour but instead must attempt to create their own internal state that will produce similar behaviour.

Replication involving reverse engineering in this way is *non-Weissmannian* (Brighton et al, 2005) in that it does not involve direct copying of a code. (Genetic replication, on the other hand, is the classic example of a Weissmannian process because DNA provides a coding scheme for phenotypes that is copied directly.) This makes linguistic structure another example of a phenotypic replicator, like memes; one that is transmitted culturally rather than biologically.

3. Transitions in linguistic evolution

It would seem, then, that language qualifies for inclusion on Maynard Smith & Szathmáry's (1995) list of major transitions for two reasons. It enables a new system of transmission of cultural content with unlimited heredity (the reason the authors themselves gave), but also its appearance heralds a completely new evolutionary system of linguistic structure itself through phenotypic replication by repeated reverse engineering.

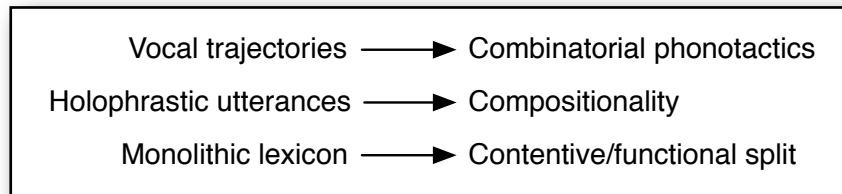
In this chapter I propose that with this insight we can take the MS&S's ideas much further and look *within* this new system of replication for evolutionary transitions that bear striking similarities to those seen in the broader evolutionary history of life. These major transitions in linguistic

¹ Croft (2000, this volume) elaborates a specific proposal along these lines, motivated by a theoretical framework for linguistic representation known as Construction Grammar. Croft treats constructions as replicators – *linguemes* in his terminology.

² Note that there are potentially confusing difference in the interpretation of “cultural”. I will use “cultural” to mean any behaviour that is inherited socially rather than genetically. “Culture” in the sense of the specific identifying practices of a society is transmitted culturally, but so too are behaviour like language.

evolution are primarily non-biological, involving an evolution of the systems of linguistic phenotypic replication, but they are likely to interact with ongoing biological evolution of humans in interesting ways that we are only beginning to understand (see also, Zuidema (2005) for discussion of transitions in linguistic evolution that have both cultural and biological implications.)

I will discuss here three candidates for major transitions:



It is important to note that these are essentially preliminary suggestions for significant shifts in the cultural evolution of language that mirror the other evolutionarily earlier major transitions. It is very likely that there will be other ones that also fit the general pattern. Equally, it is important to stress that each transition does not deliver-up the linguistic system ready for the next³ (this is also true of MS&S's list). Nevertheless, these are plausible transitions that derive from a change in the system of replication similar to that which recurs again and again in the history of life. In particular, we will see that these linguistic transitions are characterised by a loss of independent replication, and new divisions of labour – key features of the major evolutionary transitions listed in the introduction.

Before I cover each of these proposed transitions in turn, it is worth expanding a little on the process of linguistic transmission by repeated reverse engineering. In the growing literature on the cultural evolution of language, this process has been termed *iterated learning* (e.g., Kirby 2000; Kirby & Hurford 2002; Kirby, Smith & Brighton 2004; Brighton et al 2005; Kirby, Dowman & Griffiths 2007; Kalish et al 2007;):

Iterated Learning: “a process whereby some behaviour is acquired by an individual by observing an equivalent behaviour in another individual who acquired it in the same way.” (Kirby, Cornish & Smith, 2008:10681)

Language persists through iterated learning via two states: an internal state private to the individual (i.e. a grammar), and a public external state in the form of utterances (see Hurford 1990 for extensive discussion of the implications of this). Language is transformed from its internal to external state through a process of language production, and back again by perception and learning – the reverse engineering process discussed above. Note that there are many ways we could view “learning” here. At one end of the scale it could involve the acquisition of completely new structures, lexical items and so on, whereas at the other it could simply involve the adjustment of frequencies of variants driven by something like priming (see, e.g., Croft 2000; Rosenbach & Jäger 2008; Garrod & Pickering, this volume).

We normally think of this cycle of repeated production and perception/learning as driving language *change* (see, e.g., Croft 2000; Briscoe 2000; Kirby 1999; Andersen 1973), but typically iterated learning is used in models of language *evolution*. What, if any, is the difference between lan-

³ See Jackendoff (1999) for a more detailed proposal of a sequence of trajectories language could have gone through. Jackendoff's trajectories may be broadly compatible with the view presented here, but note that the driving mechanism I propose is cultural rather than biological evolution.

guage change and the cultural evolution of language, except that they are discussed in different conferences? Some authors have suggested that the term “language evolution” should be reserved for the biological evolution of the language faculty (Bickerton 2007), whereas others have used the term “evolutionary” to refer to the ongoing process of historical language change (e.g., Blevins 2004). This is basically a matter of terminology, so we should not get too hung up on it, but I think there is a useful contrast that can be made in the light of the present argument: language evolution (in the cultural, rather than biological, sense) involves language change *but also major transitions*.

It is appropriate to talk about language evolution when we are discussing the emergence of language as we recognise it today out of qualitatively different non-linguistic precursors. For some, such qualitative changes in language *require* biological changes, and hence it makes no sense to talk about cultural evolution of language. But if iterated learning leads to changes in the process of replication as I will suggest, then talking about this in terms of “mere” change fails I think to capture what is being claimed, even though the basic mechanism underlying both change and evolution are the same. To put it another way, one of the central dogmas of linguistics has been the uniformitarianist assumption: that all languages are in some sense qualitatively equivalent and have always been so (although see Newmeyer 2002 for critical review). This assumption of uniformity applies in the main to theories of diachronic linguistics, but it can hardly make sense when considering the evolutionary origins of language! Instead, it makes sense to maintain as far as possible an assumption of uniformity of *process* rather than *state*. In other words, the same process of phenotypic replication through iterated learning may drive both evolutionary transitions of language from one qualitatively different state to the next and at the same time the normal process of ongoing change within each of those evolutionary stages.

4. Three example transitions

With these preliminaries out of the way, we will now turn to the three candidates for major transitions in the evolution of language. None of these are new proposals, and the first two in particular have been described in substantial detail elsewhere as being driven by a process of iterated learning. My purpose in highlighting them here is to show that they can be viewed in the light of Maynard Smith & Szathmáry’s (1995) work as instances of a more general pattern of evolutionary transitions.

4.1. The emergence of compositionality

A number of authors have proposed that an evolutionarily early form of language may have once existed with some of the features of language as we know it, but not all (Bickerton 1990; Wray 1998; Jackendoff 1999). There are various suggestions about what such a *protolanguage* may have looked like, but one particularly prominent theory is that it consisted initially of solely *holophrastic* utterances (Wray 1998; Kirby 2000; Arbib 2005; Mithen 2005). In a holophrastic utterance, the mapping between meanings and signals is non-compositional, with the whole string of phonemes in a sentence corresponding to the whole meaning. In this respect, protolanguage utter-

ances would have been like non-compositional forms in language today such as one-word utterances (*fire!*), idioms (*bought the farm*), and holistic formulae (*how's it going?*).⁴

An obvious issue is how language evolved from this holophrastic stage into a compositional one that we are now familiar with where the meaning of most sentences is a function of the meanings of parts of the sentence. Wray (1998) proposes a process of “fractionation” by analysis where chance correspondences between sets of holistic expressions lead to learners breaking the previously unanalysed strings into sub-parts that can then go on to be combined. Although the plausibility of this process has been questioned by some (Tallerman 2007), it gains strong support from a large range of computational simulations of iterated learning (see, e.g., Brighton et al 2005, for review) and, more recently, experimental studies of the cultural transmission of artificial languages in human subjects (Kirby et al 2008). These simulations and experiments were not set up to explicitly test Wray’s (1998) hypothesis, and were developed largely independently, but nevertheless the observed process of gradual emerging fractionation by analysis is strikingly similar to that predicted by Wray (1998).

A typical simulation uses a population of agents (i.e. simulated individuals) that each learn a language mapping strings of arbitrary symbols onto a structured “world” of meanings. Agents produce a finite subset of the possible utterances in their language by being prompted with a randomly picked set of meanings drawn from the world. The set of pairs of meanings and strings thus produced are what provide the training input for the next generation of agents. Obviously, within this broad characterisation there is a lot of individual variation in different studies, for example with respect to the learning algorithm the agents employ, the space of meanings in the world, the size and structure of the population, the existence of population turnover and so on. The sentences below are examples taken from a fairly typical simulation (Kirby et al 2004): the first set are produced by an agent early in the simulation, whereas the second are from an agent thousands of generations later, after the language has evolved (purely culturally) through repeated cycles of learning and production (glosses of the meaning of the sentences are given in English):

- 1 a. ldg
“Mary admires John”
- b. xkq
“Mary loves John”
- c. gj
“Mary admires Gavin”
- d. axk
“John admires Gavin”
- e. gb
“John knows that Mary knows that John admires Gavin”

⁴ Idioms and holistic formulae are multi-word, so may appear to be compositional rather than holophrastic. However, in an idiom like *bought the farm* the meaning (*died*) has no systematic relationship to the meaning of the part of the idiom so it is non-compositional. In the case of holistic formulae such as *how's it going?* although the meaning is arguably compositional, it is likely that in actual use this utterance is not processed by either speaker or hearer compositionally, but rather as a holistic chunk. See Wray (1998) for more discussion. These apparently compositional holophrases exist in fully modern human language because it is indeed largely compositional. In the hypothesised protolanguage stage, the holophrases would have been entirely idiosyncratic in form, rather like the mono-morphemic lexicon of any modern language.

- 2 a. gj h f tej m
 John Mary admires
 “Mary admires John”
- b. gj h f tej wp
 John Mary loves
 “Mary loves John”
- c. gj qp f tej m
 Gavin Mary admires
 “Mary admires Gavin”
- d. gj qp f h m
 Gavin John admires
 “John admires Gavin”
- e. i h u i tej u gj qp f h m
 John knows Mary knows Gavin John admires
 “John knows that Mary knows that John admires Gavin”

In this example, we see clearly the emergence of compositional encoding of meanings from an initially randomly constructed holistic protolanguage. There are obvious advantages to a compositional scheme – most notably, individuals can generalise from a sub-sample of utterances in their language to unseen examples and thus communicate reliably to other individuals about novel situations. This is impossible with a holistic protolanguage since the signal for each meaning is essentially completely independent of the others and must be learned separately by rote.

Note, however, that although the compositional language is clearly superior from a communicative point of view and would therefore appear to be the sort of thing that would arise by a process of selective adaptation for communicative function, this is *not* what is driving the process in the simulation. There is no biological evolution in these simulation (nor, of course, in the parallel experimental models using human participants by Kirby et al 2008!). So what is going on here? Why do compositional languages evolve?

Rather than think about this evolution in terms of the benefits to the communicating individuals, it actually makes more sense to consider the benefits of the emergence of compositionality for the system that is being transmitted – in other words, language itself. In these models, as in reality, language persists despite having to be passed through a transmission *bottleneck*. Learners never get perfect training data because of the finite time available to them. In the simulations this is normally modelled by deliberately restricting the number of training examples so that it is lower than the total number of meanings. In the case of real language, this will always be the case because the set of meanings we may wish to convey is unbounded. Holistic languages *cannot* be reliably transmitted in the presence of a bottleneck like this, since generalisation to unseen examples cannot be reliable. A compositional language, on the other hand, enables completely reliable transmission in the presence of a bottleneck (assuming that enough training examples are available to cover the basic vocabulary of the language).⁵

⁵ Another way of thinking about this is that the bottleneck sets the level of granularity of dividing up the meanings at which a language can afford to be idiosyncratic and non-compositional. Simulation results demonstrate that frequently expressed meanings can be described holistically, whereas infrequent ones must be compositional (Kirby 2001). Similarly, languages typically exhibit irregular morphology among frequent items (Francis & Kucera 1982). Consider the past tenses of frequent verbs in English such as *be/was*, *go/went* etc. for example. In this view, the word or morpheme is simply the level at which the language can afford to be holistic and still be transmitted faithfully.

So, the language appears to have adapted simply through the process of iterated learning in such a way as to become more learnable. In the presence of a bottleneck, the structure of the strings in the language has changed in order to leverage pre-existing structure in meanings.⁶ In the simulations we can observe exactly how this is happening. It turns out to be precisely the mechanism Wray (1998) suggested: similarities between strings that by chance correspond to similarities between their associated meanings are being picked-up by the learning algorithms that are sensitive to such substructure. Even if the occurrences of such correspondences are rare, they are amplified by the iterated learning process. A holistic mapping between a single meaning and a single string will only be transmitted if that particular meaning is observed by a learner. A mapping between a sub-part of a meaning and a sub-string on the other hand will be provided with an opportunity for transmission every time *any meaning is observed that shares that sub-part*. Because of this differential in the chance of successful transmission, these compositional correspondences tend to snowball until the entire language consists of an interlocking system of regularities.

If we now return to Szathmáry's (2000) characterisation of replicators, it becomes clear that there are striking parallels between this transition from protolinguistic holophrasis to syntactic compositionality and the other evolutionary transitions discussed. In the early phase, there is a population of individual replicators corresponding to meaning-signal pairs. The survival of each of these replicators depends on it being heard by a learner. In Szathmáry's terms, these are *solitary* phenotypic replicators. To put it simply, each stands or falls on its own, they do not interact in any significant way. (See figure 1a.)

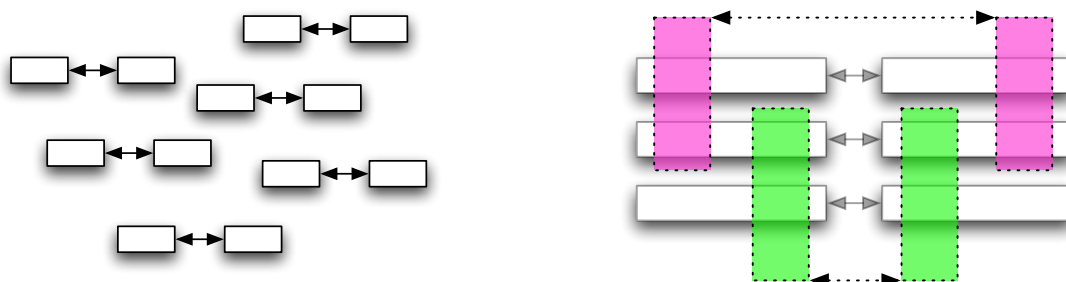


Figure 1 (a & b): Before and after the transition from a holophrastic to compositional language. Prior to the transition, replicators are individual meaning, signal pairs. After the transition, these pairs are analysable into recurring sub-pairings (e.g. words, morphemes etc.), which become the new replicators.

After the transition, sub-parts of meanings are associated with sub-parts of strings (figure 1b). The replicators now are no longer single utterances, but rather individual words/morphemes and systems of combination (i.e. grammatical rules).⁷ As we have noted, words and rules are more profligate than single holophrases, since each occurs over many meanings, but on the other hand they now rely on more examples before they can be acquired. More importantly, there is a loss of

⁶ Although, see Kirby (2007) for a model where the structure of meanings themselves can similarly be seen to adapt to improve learnability.

⁷ There is no particular need here to take a position on exactly how best to represent these replicators formally. For example, in the simulation model discussed in this section, vocabulary items and rules were represented formally using a traditional grammar formalism from computational linguistics (a DCG). Other modellers have formalised construction grammar (e.g. Steels, forthcoming), which lends itself well to an analysis in terms of replicators.

independent replicability – one of the identifying features of MS&S’s major transitions. They have become *ensemble* replicators. Along with this development comes another feature typical of an evolutionary transition: division of labour. Whereas in the holistic language, there is no distinction between a lexicon of words and the set of rules for grammatical utterances, after the transition there is a clear division between the words of the language and the systems for combining those words in grammatical or ungrammatical ways.

Although I have only discussed compositionality here, it is possible that the same process can deliver further linguistically relevant generalisations at higher levels. For example, we might consider that the major syntactic categories such as noun, verb and so on are actually generalisations over words and are thus replicators that have even greater chance of being copied once the whole ensemble of other parts of the system is in place. Once major categories are in place, then we can imagine generalisations over these categories will produce the next level of replicator. The fact that, cross-linguistically, the word-order of languages tends to place non-branching categories consistently to the left or to the right across categories (Dryer 1992) is just the kind of higher-level generalisation we might expect (see figure 2). Finally, simulations like Batali’s (2002) suggest that even basic features of language like recursion may be a side-effect of the evolution of linguistic replicators, although much more work on this is needed.

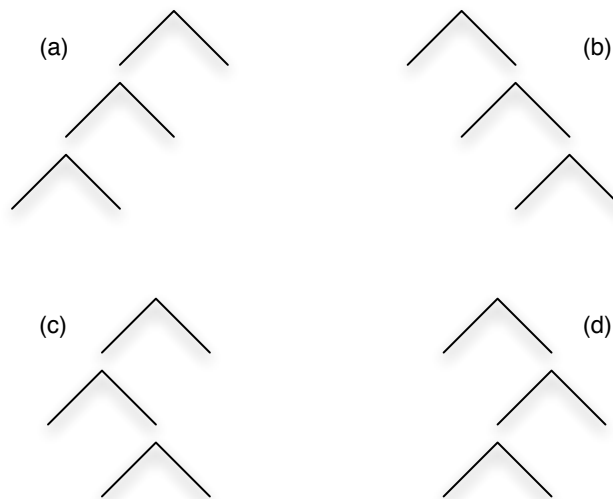


Figure 2: Branching consistency. Languages vary in the order of categories in a sentence. For example, verbs may precede or follow objects, adpositions may precede or follow noun-phrases and so on. In general this will lead to different patterns of branching when a sentence’s surface syntax is shown diagrammatically as a tree. It turns out that languages in which trees branch consistently on the left (a) or consistently on the right (b) are more common than those that branch inconsistently across the syntactic categories of a language (c & d). In other words, generalisations about the structure of a language can be made above the level of the syntactic category.

4.2. The emergence of combinatorial phonotactics

Despite the exciting results from models of the emergence of compositionality, there are some remaining concerns. For example, one question is how plausible is the reanalysis of signals? For this to happen, the protolinguistic utterances must have had analysable internal structure, otherwise no fractionation would have been possible. On the other hand, signals must not have

been too complex otherwise chance alignments that drive the process would not have occurred. In other words, protolanguage would need to have had segmentable, but non-random signals.

The next candidate for a major transition in the evolution of language delivers just such a set of signals that could feed into the one we discussed in the last section. Oudeyer (2005a; 2005b; 2006; this volume) proposes a mechanism for the emergence of combinatorial phonotactics using a computational model of interacting agents. In this section, I will summarise his model very briefly – for more details please consult the original references, particularly Oudeyer (2006) which presents many interesting simulations and has insightful discussion into their significance.

In Oudeyer's (2006) initial stage, vocalisations of the agents are arbitrary trajectories through acoustic/articulatory space. The replicators in this initial stage are these vocalisations. In the simplest version of the model, vocalisations are simple linear trajectories between a start point and an end point in a one dimensional space. One way of thinking about these trajectories is as a continuous signal moving from one part of the vocal tract to another, for example, but it is important to realise that this is really an idealisation for the purposes of understanding the general process, and Oudeyer (2006) demonstrates that this simple model can be scaled up to more complex and realistic representations of acoustic/articulatory trajectories.

Agents store representations of these vocalisations in a neural network, which is set up in such a way that start and end points of trajectories can gradually adapt to vocalisations that an agent perceives (by shifting the receptive fields of neurons towards heard vocalisations and pruning neurons that bridge between start and end points if they are rarely activated). In the simulations, a population of agents located randomly in a spatial environment are initialised with different random neural networks. At random times, an agent will produce one of the vocalisations in its repertoire at random, and then both it and the nearest other agent will adjust their neural networks in response to hearing that vocalisation.

In this way, vocalisations are culturally transmitted from one agent to another (and back again) through gradual modifications of the population's neural networks responding to vocal behaviour. This then, is another instance of iterated learning. In terms of the initial population of replicators, the survival of a vocalisation will depend on hearers tuning their neural networks to that specific vocalisation, and then going on to produce that vocalisation in turn.

So, what happens over time in this system? Figure 3a shows the repertoire of vocalisations of a typical agent at the beginning of the simulation. Vocalisations here have arbitrary start and end points. The space of vocalisations is essentially continuous, and there are no relationships between vocalisations. Rather like the case of the holistic language in the last section, these are solitary phenotypic replicators.

After repeated interactions between agents in the population, the situation looks very different however. In figure 3b, the space of vocalisations has been discretised, with trajectories only starting or ending in particular places. Furthermore, only certain *combinations* of start and end are possible. So, for example, a trajectory ending around 0.2 must start at either 0.5 or around 0.9, whereas a trajectory ending at 0.5 may start from a range of different places.

Arguably, whereas the replicators previously were individual whole continuous trajectories, now the replicators are *phonemes* and *phonotactic constraints*. Just like the words and rules of the previous section, these can no longer replicate independently. The discretised set of possible start and end points for vocalisations forms a mutually-reinforcing system of ensemble replicators. The similarity across vocalisations stabilises the neural representations of the population of agents because it is far easier for a replicator to survive that specifies a particular point as being a valid start or end of a vocalisation than one that specifies a particular trajectory through the

space. The former is reinforced by a large *set* of vocalisations produced by an agent that possesses it, whereas the latter is only reinforced by the particular vocalisation it encodes.

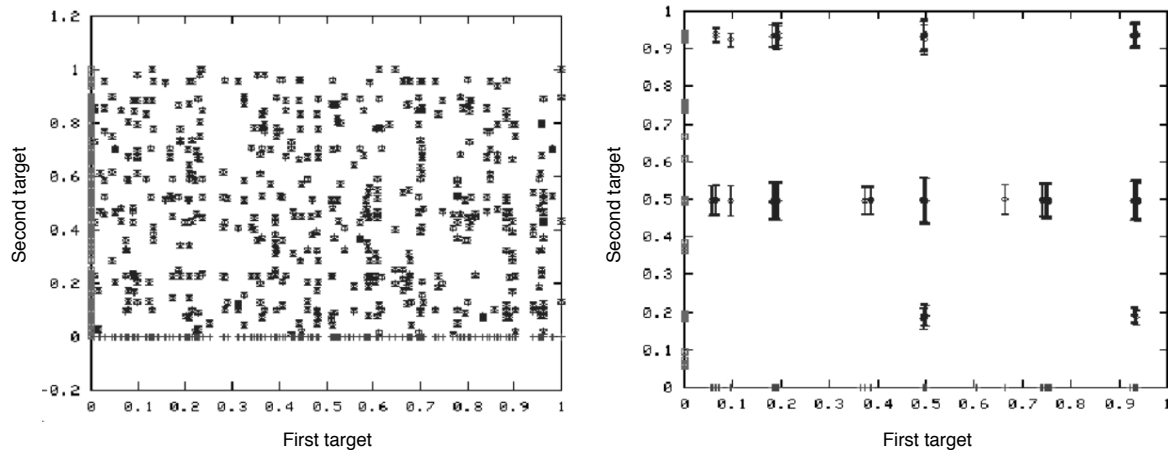


Figure 3 (a & b): Before and after the transition from a set of vocal trajectories to a repertoire exhibiting combinatorial phonotactics. The graphs show results early and late in a simulation. The two axes are start and end points of a vocalisation on an arbitrary continuum. Points plotted within the graph show possible combinations represented in a neural network architecture. See Oudeyer (2006), from where these graphs are taken, for more information.

Once again, we see a typical transition involving loss of independent replicability and the emergence of division of labour by the evolution of a new form of replicator that introduces a higher level of linguistic structure. Interestingly, this transition also shows another change in replicator type that Szathmary (2000) points out is typical in evolution: the change from analog to digital replicators.⁸

4.3. The emergence of a contentive/functional split

Given the right conditions (e.g. learning machinery, social population and so on), the basic dynamics of iterated learning seem to lead inevitably from arbitrary, continuous, vocalisations to segmentable non-random signals. Once paired with structured meanings, these provide the right kind of input to a further process of evolution by iterated learning of compositional syntax out of a holophrastic protolanguage, and then potentially on to major syntactic categories, word order universals and recursion.

It would be a mistake, however, to think the output of the second transition – to compositionality – delivers up the complexity of syntax that we see in human language. If anything, what this gives us is another kind of protolanguage. One more like fully-modern language than the holistic protolanguage of Wray (1998), but still missing much of what characterises the syntax of human language.

One of the striking differences between the kinds of languages that arise in the computational models of the emergence of compositionality, and real human language lies in the nature of the lexicon. Although we see major categories arising in the simulation models these are largely semantically motivated. In many ways the lexicon that emerges is homogeneous in its role in the

⁸ Szathmary (2000) uses the terms *holistic* or *processive* for the former, and *modular* for the latter, but this terminology may be confusing given the different use of the term “holistic” in the discussion of Wray’s (1998) protolanguage.

communication system: each word maps onto part of the content of an utterance and combines with other words to compose a whole message.

Obviously there are a large number of lexical items in real languages that behave in this way. These *contentive* elements, such as nouns, verbs, adjectives and so on form the *open-class* subset of any language's lexicon. However, in language there is another class of lexical items – the *functional* elements, such as determiners, prepositions, markers of case, aspect, tense, number and so on, which form the *closed-class* subset of the lexicon.

This divide between contentive and functional is perhaps the most fundamental in language. Contentive elements are open-ended, phonologically diverse, and meaningful, whereas functional elements are closed-class, have atypical phonology, and serve grammatical roles. There is evidence that these two sets of lexical items are processed differently in the brain, and the distinction between them has an important role in the early acquisition of language (see, Cann 2001 for review of the various aspects of this division).

Functional elements in particular are a crucial part of the syntax of fully-modern human language. They underpin much of what makes the structure of language so interesting and unusual, and in some theories of grammar they have a basic role to play in defining the particular structure of a language and parameterising cross-linguistic variation (Cann 2001).

Arguably, the change from a monolithic lexicon to one with a basic split between contentive and functional is the last major transition that leads us to language as we know it today (see figure 4). Once again, this fundamental split in the lexicon reflects a very common feature of the evolutionary transitions: division of labour. Replicators in an ensemble will over time become increasingly co-dependent and tend to diversify in function. But what drives this transition in language?



Figure 4 (a & b): Before and after the transition from a monolithic lexicon to a contentive/functional split. Prior to the transition, the lexicon contained contentive elements only distinguished by broadly semantically defined syntactic categories. Afterwards, the lexicon specialises into an open class of elements carrying content, and a closed class of those with a grammatical function.

At this stage, we do not have a model that is as simple as the ones described above for the other transitions, but some computational modellers are now turning to the problem of the features like case-coding (which, when marked overtly, involves closed-class morphemes) and how it may emerge from the dynamics of socially interacting communicative agents (e.g., Steels, forthcoming).

Furthermore, it is likely we can learn much from studies of the ongoing emergence in languages of functional elements out of previously contentive items through the process of *grammaticalisation* (e.g., Hopper & Traugott 1993). Ultimately, grammaticalisation is driven by iter-

ated learning in that it is a product of the cultural transmission of language. So, we can say with some confidence that the phenotypic replication of language can lead to the emergence of new replicators with a specifically syntactic function because we can actually observe this process in languages today. What remains an open question at this stage is if we can demonstrate that the same process lead to the emergence of the very first functional elements in the longer term emergence of language. In other words, did the process of grammaticalisation lead to the original split in the lexicon in the first place? The answer to this question must await further modelling.

5. Summary and conclusions

In this paper, I have tried to develop parallels between ongoing research into the cultural evolution of language and the generalisations made by Maynard Smith & Szathmáry (1995) about transitions in evolutionary systems more broadly. I have argued that ongoing change in language is driven by phenotypic replication through iterated learning, and that this same process also delivers major transitions in the emergence of language which justify us calling this non-biological process “evolution” rather than merely “change”.

The transitions I have surveyed all involve the emergence of key aspects of linguistic structure out of a previous state in which that structure was not present. They have suggestive similarities with each other, and with evolutionary transitions in general. In particular, they involve a shift (at various levels) from solitary replicators with independent replicability, to ensemble replicators that lose independence and lead to diversification and division of labour as a result.

Although these transitions are purely cultural, in the sense that they can arise out of the process of repeated transmission of language through production, observation, and learning even among biologically identical agents, this does not mean that there is no role for biological evolution. For example, we need to consider what the conditions must be in order for iterated learning to happen at all. Species-specific social and cognitive adaptations must be involved to make us such avid transmitters of linguistic replicators. Secondly, the emergence and ongoing evolution of a structured linguistic system is likely to have a knock-on effect on the biological evolution of its carriers (us). It is entirely reasonable to expect co-evolution of the culturally transmitted linguistic system and the biologically determined apparatus for supporting it (see, e.g., Smith & Kirby 2008; Dediu, this volume; Deacon, this volume, for discussion).

Maynard Smith & Szathmáry’s (1995) work provides a rich framework for thinking about replication. They themselves identified the importance of language in this light, but language is a new system of replication in more than one sense: it is both an enabler of cultural replicators with unlimited heredity, and also a new kind of evolutionary system itself. Iterated learning is the process of linguistic transmission, and it drives both language change and the transitions to qualitatively new kinds of linguistic system. By seeing language as an evolutionary system, the biggest payoff we get may be the ability to take biologists’ insights into the evolution of life and apply them to the evolution of language.

6. Bibliography

- Andersen, H. (1973). Abductive and deductive change. *Language*, 40:765–793.
- Arbib, M. (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences*, 28:105–167.
- 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.

- Bickerton, D. (1990). *Language and Species*. University of Chicago Press.
- Bickerton, D. (2007). Language evolution: a brief guide for linguists. *Lingua*, 117(3):510–526.
- Blevins, J. (2004). *The Evolution of Phonology*. Cambridge University Press.
- Brighton, H., Smith, K., and Kirby, S. (2005). Language as an evolutionary system. *Physics of Life Reviews*, 2:177–226.
- Briscoe, E. J. (2000). Grammatical acquisition: Inductive bias and coevolution of language and the language acquisition device. *Language*, 76(2):245–296.
- Cann, R. (2001). Functional versus lexical: A cognitive dichotomy. In Borsley, R. D., editor, *The Nature and Function of Syntactic Categories*. Academic Press.
- Croft, W. (2000). *Explaining Language Change: an evolutionary approach*. Longman.
- Dryer, M. (1992). The Greenbergian Word Order Correlations. *Language* 68: 81-138
- Francis, N. and Kucera, H. (1982). *Frequency Analysis of English Usage: Lexicon and Grammar*. Houghton Mifflin.
- Hopper, P. J. and E. C. Traugott (1993) *Grammaticalization*. Cambridge, Cambridge University Press.
- Hurford, J. (1990). Nativist and functional explanations in language acquisition. In Roca, I. M., editor, *Logical Issues in Language Acquisition*, pages 85–136. Foris, Dordrecht.
- Jackendoff, R. (1999). Possible stages in the evolution of the language capacity. *Trends in Cognitive Sciences*, 3(7):272–279.
- Kalish, M. L., Griffiths, T. L., and 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*. Oxford: OUP.
- 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 Transactions on Evolutionary Computation*, 5(2):102–110.
- Kirby, S. (2007). The evolution of meaning-space structure through iterated learning. In Lyon, C., Nehaniv, C., and Cangelosi, A., editors, *Emergence of Communication and Language*, pages 253–268. Springer Verlag.
- Kirby, S., Cornish, H., and 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(31):10681–10686.
- Kirby, S., Dowman, M., and Griffiths, T. (2007). Innateness and culture in the evolution of language. *Proceedings of the National Academy of Sciences*, 104(12):5241–5245.

- 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*, chapter 6, pages 121–148. Springer Verlag, London.
- Kirby, S., Smith, K., and Brighton, H. (2004). From ug to universals: linguistic adaptation through iterated learning. *Studies in Language*, 28(3):587–607.
- Maynard-Smith, J. and Szathmáry, E. (1995). *The Major Transitions in Evolution*. New York: Oxford University Press.
- McMahon, A.M.S. (1994). *Understanding language change*. Cambridge: CUP.
- Mithen, S. (2005) *The Singing Neanderthals: The Origins of Music, Language, Mind and Body*. Weidenfeld & Nicolson.
- Newmeyer, F. J. (2002) Uniformitarian Assumptions and Language Evolution Research. In Alison Wray, editor, *The Transition to Language*. Oxford: Oxford University Press.
- Oudeyer, P.-Y. (2005a). The self-organization of combinatoriality and phonotactics in vocalization systems. *Connection Science*, 17(3-4):325–341.
- Oudeyer, P.-Y. (2005b). From holistic to discrete speech sounds: The blind snow-flake maker hypothesis. In Tallerman, M., editor, *Language Origins: Perspectives on Evolution*, chapter 4, pages 68–99. Oxford University Press.
- Oudeyer, P.-Y. (2006). *Self-organization in the Evolution of Speech*. Oxford University Press.
- Pinker, S. and Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences*, 13(4):707–784.
- Ritt, N. (2004). *Selfish sounds and linguistic evolution: A Darwinian approach to language change*. Cambridge: CUP.
- Rosenbach, A. & Jäger, G. (2008) Priming and unidirectional language change, *Theoretical Linguistics* 34(2), 85-113.
- Smith, K. and Kirby, S. (2008) Cultural evolution: implications for understanding the human language faculty and its evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1509):3591--3603.
- Steels, L. (forthcoming). Cognition and social dynamics play a major role in the formation of grammar. In D. Bickerton and E. Szathmáry (eds). *Biological Foundations and Origin of Syntax*. Cambridge, MA: MIT Press.
- Szathmáry, E. (2000). The evolution of replicators. *Philosophical Transactions of the Royal Society B*, 355:1669–1676.
- Tallerman, M. (2007). Did our ancestors speak a holistic protolanguage? *Lingua*. 117(3): 579-604.
- Wray, A. (1998). Protolanguage as a holistic system for social interaction. *Language and Communication*, 18(1):47–67.
- Zuidema, W. (2005). *The Major Transitions in the Evolution of Language*. PhD Thesis, University of Edinburgh.