The adaptation of language toward increasingly faithful replication during its cultural transmission

INTRODUCTION

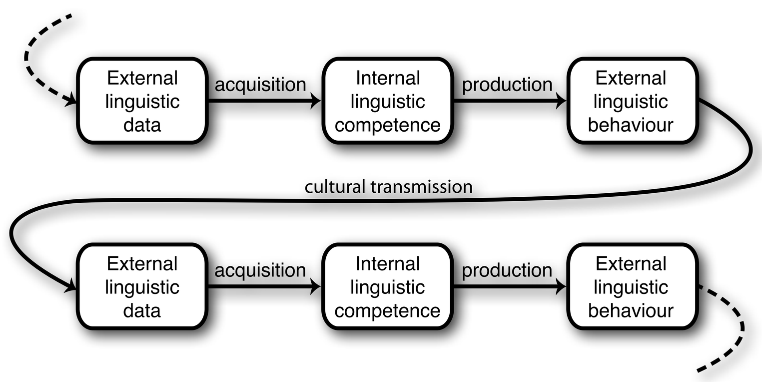
Evolutionary Linguistics is a highly interdisciplinary field bridging such subject areas as Anthropology, Archaeology, Evolutionary Biology, Psychology, and of course Linguistics. As such, there is currently little consensus on how human language emerged in our species' history. One suggestion has been that human language arose as a function of three adaptive processes: evolution by natural selection, individual learning, and cultural evolution (Kirby & Hurford, 2002). In this paradigm, biological evolution is intrinsically difficult to explore due to the ephemeral nature of language; however, the interface between learning and cultural evolution has recently been tested in a variety of mathematical, computational, and experimental models. These models demonstrate that systematic linguistic structure can emerge in the transmission of language across multiple language users. Kirby, Cornish, and Smith (2008) introduced an experimental method for studying the cumulative effect of such cultural transmission, and their iterated learning model represented the first experiment on human participants to suggest that the transmission of language leads to the appearance of compositional linguistic structure without any explicit designer.

This poster presents the theory and methodology behind this experiment, alongside the results I have obtained in running my own version of the experiment. My results showed that languages become easier to learn as they are transmitted along a line of language users; however, the emergence of compositional linguistic structure was not forthcoming. This presents a problem: if the languages evolve to become easier to learn, yet structure does not emerge, then there must exist at least one other mechanism by which the languages optimize their faithful replication. The results suggest that this mechanism might lie in a different type of adaptation: over the course of the experiment, the languages tended to become more focused on a smaller set of syllable patterns, which appears to make them easier to learn.

THEORY

The dominant explanation for how language arose in the human species has been that language conferred a useful ability on humans, and was therefore selected for in the process of evolution by natural selection (e.g. Pinker & Bloom, 1990). More recently, many scientists have questioned this view, at least in part, and have suggested that language may constitute an evolutionary system in its own right which has adapted to human cognitive mechanisms (e.g. Christiansen & Chater, 2008). This thesis is related to Richard Dawkins's concept of memetics (1976), in which units of culture (memes) adapt to their environment (human brains) as they are transmitted horizontally between members of a group, and vertically down the generations.

Consider firstly the way language is culturally transmitted: It is acquired by a child from the language that exists in his or her environment; this provides the child with internal linguistic competence. Later this individual will exhibit external linguistic behaviours which become the source of linguistic data for a new generation, and so forth ad infinitum. This is illustrated in the following diagram (adapted from Smith et al. (2003: p.540)):

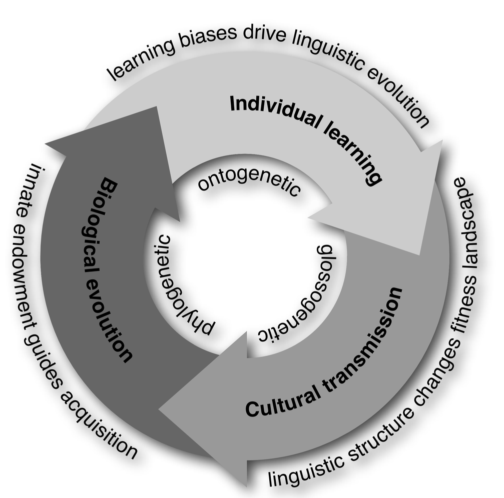


This process is known as cultural transmission, and it can be equally applied to other cultural artefacts such as tool making, agricultural practices, and religion.

The line marked cultural transmission in the diagram above suggests that the transmission of language is a clean, tidy process, but in fact transmission is prone to copying errors, and is anything but faithful. Such errors are introduced by, for example, constraints on working memory, disfluency, and noise in the environment. However, such errors are important to the development of a language because mutation introduces a source of variation, and variation leads to selection. Thus, with each passing generation of language user, the language is shaped by its environment, which may be taken as human brains and the spacial and temporal distances that must be overcome in order to maintain its successful transmission. This is illustrated here (adapted from Kirby et al. (2004: 600)):

A variety of mathematical, computational, and experimental models of these phenomena have been conducted, which have shown that unsystematic artificial 'languages' can self-organize under the pressure involved in transmission, gaining properties such as compositionality and recursion epiphenomenally (see Hurford 2002, Brighton et al. 2005, Cornish et al. 2009, Scott-Phillips & Kirby 2010 for reviews).

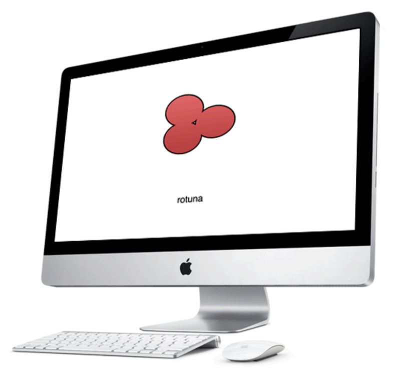
It is important to note that this thesis is not opposed to a part for natural selection in the evolution of human language; rather, it is suggested that language emerged as a product of both biological and linguistic evolution. And in fact, a third adaptive process seems to feed into this system too – individual learning. Specifically, it is believed that the way in which children learn language during their development has affected the universal features we observe in language today. The interaction between these three adaptive processes is illustrated in the diagram to the right (adapted from Kirby & Hurford, 2002: p.122): the innate leaning biases, which define how a child learns, affect the properties of his or her language; then, as the language is passed from one generation to the next, the properties of the language adapt to increase their successful transmission; this modifies the fitness landscape, and biological evolution selects for an innate endowment that defines learning biases which map closely on to the linguistic structure present in the environment; and so forth.



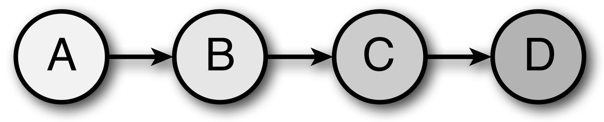
EXPERIMENTAL METHOD

Kirby et al. (2008) came up with an experimental paradigm to test this theory. In this study I replicated their methodology but introduced some new variables, including an auditory mode, a new semantic space, and a different signal space.

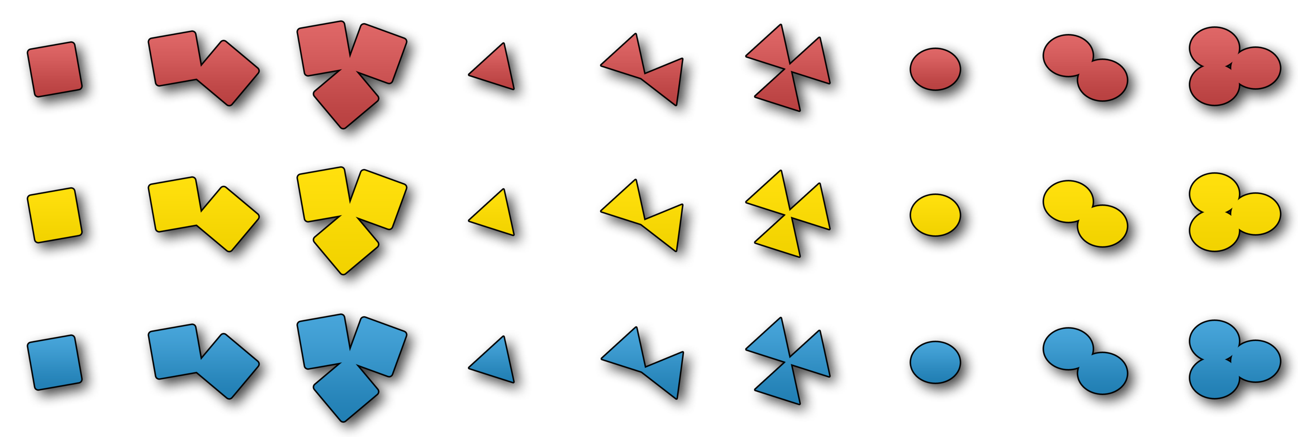
Thirty participants were recruited to take part in an 'alien language experiment'. There were no prerequisites for participation other than a minimum age of 18 years. Participants were predominantly undergraduate students. The female:male sex ratio was 16:14. The mean age was 24.5. Participants were given written and verbal instructions, which portrayed the experiment as a simple memory game. Specifically, participants were told that they would be learning words for alien fruit, and that their task was to memorize the mapping of words to images to the best of their abilities. Participants were not told about the cultural nature of the experiment, and were not primed to recognize or produce linguistic structure. Having received instructions, participants were left alone at a computer terminal to execute the experiment, which typically took around 20 minutes to complete.



Following Kirby et al. (2008), the experiment adopts a diffusion chain design, in which the output of a given participant becomes the input for the following participant. Three separate diffusion chains were run, labelled herein A–C. In each chain, an artificial 'language' was diffused down a line of ten participants, representing ten cultural generations. Each chain was initiated with a separate 27-word language, which was generated by randomly concatenating two-, three-, or four-syllable words from a syllabary of 28 syllables. Each participant was randomly assigned to one of the three chains, and acquired the language in one session.



The visual stimuli were a set of 27 simple graphics (see below), which conform to a strict 3 × 3 × 3 semantic-space, that is, the fruit come in three colours (red, yellow, blue), three shapes (square, circle, triangle), and three 'numbers' (one, two, or three segments). During three training periods, the fruit were each presented for 5 seconds alongside their associated labels, both orthographically as a string of lower case letters, and acoustically by means of a computationally synthesized voice. Each period of training was followed by a test, which prompted the participant to input the names associated with particular fruit graphics. There is one important twist, however: during training, participants were only taught 14 of the 27 fruit, but were tested on all 27 in the test. This enforces a transmission bottleneck: since participants cannot learn all 27 strings, there exists a cumulative and invisible pressure to circumvent the bottleneck by exploiting the structure inherent in the semantic-space.



STATISTICAL METHODS

According to the theory outlined to the left, the initially random languages should become increasingly easy to learn due to increasingly compositional linguistic structure. To measure these qualities, I followed Kirby et al. (2008):

Measure of transmission error The mean edit distance between the strings in a participant's output and the corresponding strings in the previous participant's output provides a measure of intergenerational transmission error, and is given by

$$E(i) = \frac{1}{|M|} \sum_{m \in M} \frac{\mathrm{LD}(s_i^m, s_{i-1}^m)}{\max(|s_i^m|, |s_{i-1}^m|)},$$

where mean transmission error E at generation i is 1 over the magnitude of meaning set M, multiplied by the sum of the normalized Levenshtein edit distances for each meaning m in meaning set M. The Levenshtein edit distance is computed between string s (for meaning m at generation i) and string s (for meaning m at generation i–1), and normalized by dividing by the length of the longer string. This measure is expressed over the interval [0,1], where 0 is no error and 1 is maximal error.

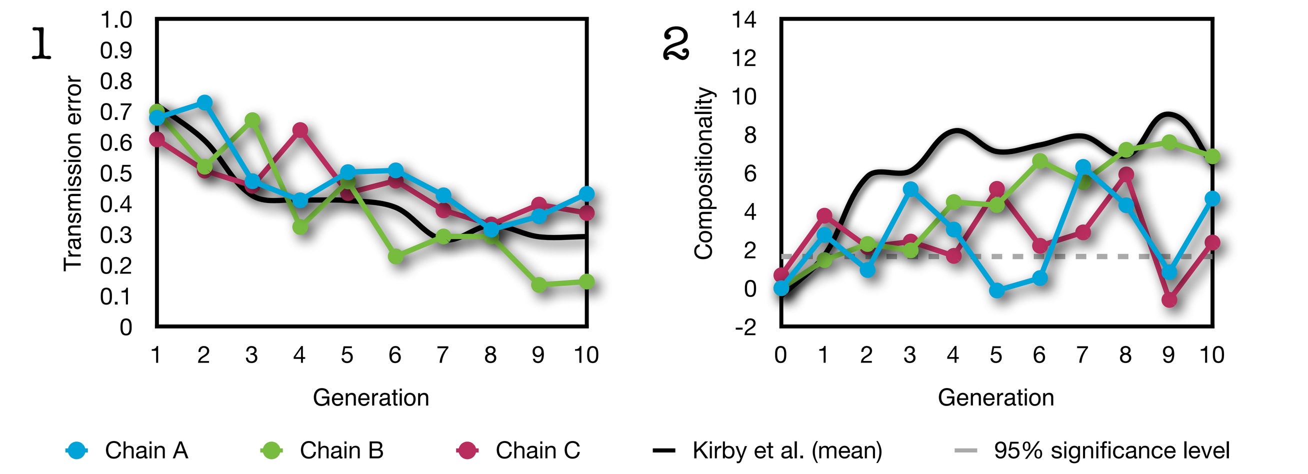
Measure of compositionality A correlation between the difference in form and the difference in meaning for each possible pairing of strings in a given participant's output provides a measure of the degree to which similar meanings are expressed by similar forms, and is calculated as follows: firstly the normalized Levenshtein edit distance is computed for each of the 351 possible pairings of strings in a participant's 27-string output; a hamming distance is then computed for each of the corresponding pairs of meanings; finally a Pearson's product-moment correlation coefficient is calculated between these edit distances and their corresponding hamming distances, giving an indication of their statistical alignment.

To measure significance and standardize the coefficients across languages, a Monte Carlo sample of 10,000 randomized permutations of each participant's output is produced, against which the participant's veridical coefficient is compared. This is performed by calculating the mean and standard deviation of the coefficients in the Monte Carlo sample, and then deriving the standard score, which constitutes our final measure of compositionality. To put this measure in context, a randomly generated language scores around 0, while English scores around 17.7.

RESULTS

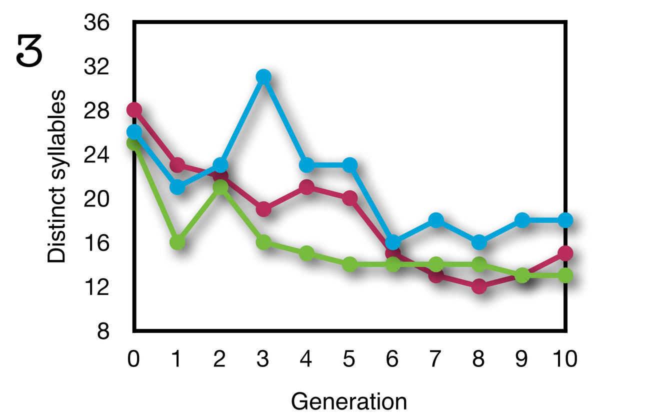
The graphs below show the experimental results for ten participants in each of three diffusion chains labelled A–C. Graph 1 shows a statistically significant decrease in transmission error between the initial and final generations (mean decrease = 0.346, SD = 0.179, t (2) = 3.358, p < .04). This is in line with the predictions and results of Kirby et al. (2008): the languages clearly adapt to increase their own transmissibility by becoming easier to learn over the course of the experiment. Furthermore, the effect is cumulative: the correlation between generation number and transmission error is statistically significant (r = -0.762, N = 30, p < .001).

The increase in learnability is a product of the languages' increasingly compositional structure, and graph 2 confirms that compositionality increased significantly between the initial and final generations (mean increase = 4.408, SD = 2.589, t (2) = 2.949, p < .05). Furthermore, the majority of participants' outputs (those above the dashed 95% significance line) are statistically nonrandom, and – as with transmission error – the increase in compositionality is cumulative: the correlation between generation number and compositionality score is statistically significant (r = 0.461, N = 33, p < .004).



We can demonstrate the compositional structure that emerged by looking more closely at the languages. The table below illustrates how the linguistic signals map on to the semantic-space in the language of one particular participant (B9). Shape is expressed by the first morpheme: pi- or do- in most cases of square fruit; pa- or po- in the case of triangular fruit; and ro- in all cases of circular fruit. Colour is not clearly expressed by any component, although final -ni tends to occur more often among red fruit, while final -na tends to occur among blue and yellow fruit (especially noticeable among the circular fruit). Number is also not clearly expressed, although medial -ni- and -ti- may mark a singular/plural dichotomy (again this is especially noticeable among the circular fruit). Note also that square fruit with three segments have a noticeably different (but internally regular) form from other square fruit which may be considered a marker of number in its own right.

	red	yellow	blue	
square	pisa	pifa	pifa	1
	pifa	fifa	pifa	2
	dorisi	dorisi	dorini	3
triangle	pani	pani	pafi	1
	panina	pofini	panina	2
	panini	pofina	pofi	3
circle	ronini	rona	ronina	1
	rotini	rotina	rotina	2
	rotini	rotina	rotina	3



Graphs 1 and 2 above also plot (in black) the average transmission error and compositionality for the languages that evolved in Kirby et al. (2008). While our results for transmission error show a very similar mean, range, and trajectory, our results for compositionality are significantly different (mean difference = 3.174, t (68) = 4.57, p < .001). This presents a problem: how can the observed decrease in transmission error be explained by such a small increase in compositionality?

I believe this problem can be solved by looking at another way in which the languages adapted. As shown in graph 3, there was a significant decrease in the variety of syllables being employed between the initial and final set of languages (mean decrease = 11, SD = 2.646, t (2) = 7.201, p < .02). Furthermore, there is a strong correlation between the number of distinct syllables in a participant's learning material and his or her transmission error (r = 0.633, N = 30, p < .001). This suggests that the decrease in transmission error might also be explained, at least in part, by the erosion of the signal-space. Since these languages can operate compositionally and expressively with as few as nine morphemes (or even fewer, if they employ a strict morphological order), and given that the languages are easier to learn if there are fewer syllables/syllable patterns to memorize, the languages gladly sacrifice signal-space to the benefit of learnability. In other words, it makes sense from a language's perspective, to relinquish extraneous syllables that only serve to confuse the learner. This kind of adaption does not occur to a significant extent in Kirby et al. (2008), since the initial languages were concatenated from a relatively small syllabary.

CONCLUSIONS

The present experiment supports the findings of Kirby et al. (2008), and validates their experimental paradigm. It shows that compositional structure can emerge from an initially unsystematic language over repeated cycles of expression and induction. Notably, it shows that the degree of structure that emerges in a language is determined by its linguistic environment, and that there is at least one other way in which languages can adapt.

From a practical angle, this paper shows that it is possible to add an auditory modality to the learning regimen of this iterated learning model, which could be particularly useful to research grounded more firmly in the study of phonetics.

The next major step in developing a fully fleshed-out theory of language evolution by cultural transmission will be to show empirically that compositionality, as well as other universal properties of language, can emerge under environments that more closely resemble the real world.

Jon William Carr
j.w.carr@sms.ed.ac.uk
Language Evolution and Computation Research Unit
School of Philosophy, Psychology and Language Sciences
University of Edinburgh, United Kingdom

REFERENCES

Brighton, H., Smith, K., & Kirby, S. (2005). Language as an evolutionary system. *Physics of Life Reviews*, *2*(3), 177—226.

Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, *31*(5), 489—558.

Cornish, H., Tamariz, M., & Kirby, S. (2009). Complex adaptive systems and the origins of adaptive structure: What experiments can tell us. *Language Learning*, *59*(s1), 187—205.

Dawkins, R. (1976). Memes: The new replicators. In *The Selfish Gene* (pp. 203—215). Oxford: Oxford University Press.

Hurford J. R. (2002). Expression/induction models of language evolution: Dimensions and issues. In T. Briscoe (Ed.), *Linguistic Evolution through Language Acquisition* (pp. 301—344). Cambridge: Cambridge University Press.

Kirby, S., & Hurford, J. R. (2002). The emergence of linguistic structure: An overview of the iterated learning model. In A. Cangelosi & D. Parisi (Eds.), *Simulating the Evolution of Language* (pp. 121—147). London: Springer Verlag.

Kirby, S., Smith, K., & Brighton, H. (2004). From UG to universals: Linguistic adaptation through iterated learning. *Studies in Language*, *28*(3), 587—607.

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 of the United States of America*, *105*(31), 10681—10686.

Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences*, *13*(4), 707—784.

Scott-Phillips, T. C., & Kirby, S. (2010). Language evolution in the laboratory. *Trends in Cognitive Sciences*, *14*(9), 411—417.

Smith, K., Brighton, H., & Kirby, S. (2003). Complex systems in language evolution: The cultural emergence of compositional structure. *Advances in Complex Systems*, *6*(4), 537—558.

This research was conducted while I was an undergraduate at the University of Glasgow. I wish to thank Wendy Anderson (University of Glasgow), Simon Kirby, Hannah Cornish (University of Edinburgh), and Scott Seyfarth (University of California, Santa Cruz) for their kind input, advice, and encouragement. Thanks must also be extended to all the participants who helpfully took part in this experiment.