The role of practice and literacy in the evolution of linguistic structure

THE ROLE OF PRACTICE AND LITERACY IN THE EVOLUTION OF LINGUISTIC STRUCTURE

MÓNICA TAMARIZ, J. ERIN BROWN & KEELIN M. MURRAY

Language Evolution and Computation, PPLS, The University of Edinburgh,
3 Charles Street, Edinburgh EH8 9AD, UK.
monica@ling.ed.ac.uk; erin@ling.ed.ac.uk; k.m.murray-3@sms.ed.ac.uk.

Recent iterated language learning studies have shown that artificial languages evolve over the generations towards regularity. This trend has been explained as a reflection of the learners' biases. We test whether this learning bias for regularity is affected by culturally acquired knowledge, specifically by familiarity and literacy. The results of non-iterated learning experiments with miniature artificial musical and spoken languages suggest that familiarity helps us learn and reproduce the signals of a language, but literacy is required for regularities to be faithfully replicated. This in turn indicates that, by modifying human learning biases, literacy may play a role in the evolution of linguistic structure.

1. Introduction

Throughout the history of our species, language has been transmitted between generations by observation of the linguistic behaviour of conspecifics and subsequent practice of the skills learned during language use. Only structures that can be transmitted in this way (those that go through the "transmission bottleneck") exist in any language (Kirby & Hurford, 2002). Linguistic structure reflects the conditions where language is used and transmitted. Structure may thus emerge from individual processes such as our ability to establish relationships between symbolic units (Deacon 1997) or from pressures derived from social communication (Schoenemann, 1999) and the cultural transmission of language (Kirby & Hurford, 2002). One suggested source of linguistic structure related to both use and transmission is literacy, a powerful cultural institution that touches on all aspects of language. Literacy enhances phonological (Ehri, 1985) and morphological (Nunes et al. 2006) awareness and seems to be necessary for the segmentation of utterances into words (Olson, 1996; Ramachandra & Karanth, 2007), for grammaticality judgments related to syntax (de Villiers & de Villiers, 1972) as well as for the ability to analyze linguistic structure separately from semantic content (Karanth & Sutchira, 1983). This paper is concerned with the idea that cultural institutions like teaching and literacy may have modified the transmission bottleneck of

language, allowed different kinds of linguistic structures to exist and thus influenced the course of language evolution.

Recent experimental approaches to cultural evolution (e.g. Kirby, Cornish & Smith, 2008; Kalish, Griffiths & Lewandowsky, 2007) show how culturally learned and transmitted information comes to reflect the inductive biases of the learners. The biases at work in the above-mentioned studies include a preference for linear functions and a preference for regular mappings between signals and meanings in artificial miniature languages. In the case of the languages, the fact that the products of repeated transmission share some of the characteristics of language indicate that the evolving systems have successfully adapted to the learners' expectation of what a language looks like. This expectation may be influenced by innate and culturally acquired biases. The present study explores the role of culturally acquired knowledge on individual learning of language structure. Specifically, we look at the effect of literacy and of extensive experience of language on how the *structure* of artificial miniature language is learned and reproduced. We hypothesize that bias for regular mappings observed in recent studies (e.g. Tamariz & Smith, 2008; Kirby, Cornish & Smith, 2008) is enhanced by literacy as well as by familiarity with language forms and functions.

2. A musical artificial language experiment

In our investigation we turned to musical literacy as a proxy for orthographic literacy, since finding two groups of participants matched in all aspects except literacy proved impossible. Participants were shown artificial miniature musical (or spoken) systems consisting of mappings between tunes (or pseudo-words) and drawings and subsequently tested how well they had learned the systems. We manipulated two independent variables: participant literacy and level of regularity of the input languages.

2.1. *Participants*

85 undergraduate students were recruited through the university employment website. They fell into one of four categories: 1. Musician participants who had studied music at the University level beyond year one, could read musical notation and play an instrument proficiently and were currently practicing (the mean length of practice was 15.2 years). 2. Illiterate musicians who were required to play an instrument and practice regularly, but to have had no instruction in musical notation. This condition allows us to isolate the effects of literacy from those of extensive practice and familiarity with music. 3. Non-

musicians, who were required to have no musical background. 4. A fourth group of participants with no musical requirements was assigned to the Spoken language condition. Groups 1 and 4 were literate in their respective artificial languages, but, while spoken language usually has referential meaning, music usually has not. Comparing these two conditions may tell us something about the effect of habit of referentially linking the signals in the (musical or spoken) language with meaning

2.2. *Materials*

Three musical and three spoken languages each comprising 27 meaning-signal pairs were used in the experiment. The meanings used all languages 27 were visually presented figures (Tamariz & Smith, 2008) including all possible combinations of three shapes (square, circle, hexagon), three colours (red, blue, yellow) and inserts (star, dot, cross).
The musical signals were constructed by combining two-note intervals. The nine intervals used were perceptually distinct and positioned within the human voice range. All intervals had middle C (*C4,* 262 Hz) as the first note, followed by one of nine other notes (see Fig. 1).

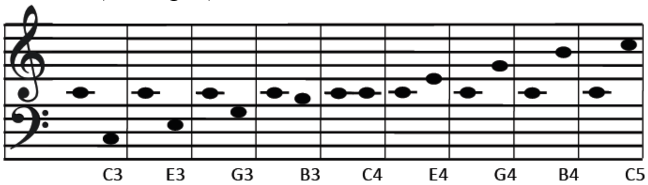


Figure 1. The nine signal units used to construct the signals for the musical languages. All two-note intervals begin with C4.

Each note in the interval was played as a pure tone for 330 ms and there were 150 ms of silence between them. Signals were constructed by concatenating three intervals, with 550 ms of silence between intervals. The spoken signals were the following nine consonant-vowel syllables: pe, mu, lo, tu, na, di, be, ga, ki.

Three languages (mappings of the above signals and meanings) were constructed. In the fully compositional language *L3*, color is regularly encoded in the first musical interval, shape in the second and insert in the third (e.g. figures of the same colour start with the same interval, and figures with the same insert have the same final interval). The partially compositional language *L2* generally follows the same pattern, but has some exceptions or irregularities in each meaning dimension. Random language *L1* contains no regular mappings. All materials were presented via a *Psyscope* script run on an Apple Macbook.

Participants were only exposed to one half of the signal-meaning items in the relevant language (L1, L2 or L3), as a 50% transmission bottleneck was in place. (They were, however, tested on all 27 meanings.)

2.3. *Data analysis*

We analyzed the fidelity of reproduction of the signals and the fidelity of reproduction of the compositional structure in the output languages. The *reproduction of the signals* was the number of times that a segment in the input language was reproduced exactly, in the same position for the same signal, in the output language. (Note that only signals in the items that were seen (those left after applying the bottleneck) were counted.) The *reproduction of the compositional structure* of the input language in the output language was quantified using *RegMap* (Tamariz, *in press*; Cornish, Tamariz & Kirby, *in press*), an information-theoretical tool designed to quantify the regularity of the mappings between two domains. The compositional patterns in the languages were quantified with partial *RegMap*. Partial *RegMap* is calculated for each meaning component-signal component pair (e.g. for colour and initial segment). The two conditional entropies *H(S|M)* and *H(M|S)* (Eqn. 1) of the cooccurrence frequency matrix between each meaning component variant *M* (e.g. red, blue, yellow) and signal component variant *S* (e.g. *be, ga, ki*) in a language are obtained. These are used in equations 2 and 3 to obtain *RegMap(S|M)* and *RegMap(M|S)*, which are combined in equation 4 to yield the corresponding partial *RegMap*. A partial *RegMap* value measures the confidence that a meaning component is reliably and unambiguously associated with a signal component, given a learner's experience of a language (reflected in the cooccurrence frequency matrix).

(1) $$H(X \mid Y) = -\sum_x \sum_y p(y) p(x \mid y) \log_2(p(x \mid y))$$

(2) $$\mathrm{Re}gMap(S \mid M) = 1 - \frac{H(S \mid M)}{\log(n_s)}$$

(3) $$\mathrm{Re}gMap(M \mid S) = 1 - \frac{H(M \mid S)}{\log(n_m)};$$

(4) $$\mathrm{Re}gMap(S, M) = \sqrt{\mathrm{Re}gMap(S \mid M) \times \mathrm{Re}gMap(M \mid S)}$$

Reproduction of the compositional structure of a language by a participant is the correlation (Pearson's *r*) between the partial *RegMap* values of the participant's input and output languages. This measure captures reproduction of the compositional structure independently of reproduction of the signal elements. For example, if the output of a participant exposed to language L3

maps colour to the first segment, shape to the second and insert to the third, the correlation will be high even if the actual segments he or she produces are different from those in the input language.

2.4. *Procedure*

Experiments were conducted individually and took approximately 30 to 45 minutes. 85 participants were run in total. They were told they would learn a language before they underwent three *training* phases and one testing phase, with breaks between phases. Each training exposure consisted of visual presentation of a meaning and, 50 msec later, auditory presentation of the corresponding musical or spoken signal. The signal was played through headphones three times, leaving silence (6 seconds for musical languages, 4.5 sec for spoken language) after each signal where participants were instructed to repeat it. During the *testing* phase, meanings were visually presented one at a time and the participants had to hum or say the corresponding signals. Participants' productions were recorded and later analyzed using Adobe Audition 1.5, which assigned the frequency of each note to the appropriate semitone. This analysis was checked by the experimenters.

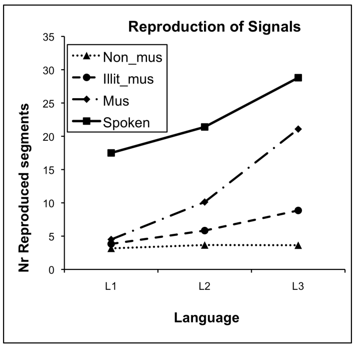
2.5. *Results and discussion*

The levels of reproduction of the signals (Fig. 2a) reflect the degree to which participants learned and faithfully reproduced the signal segments and their positions during training; reproduction of the compositional structure (Fig. 2b) reflects the degree to which participants learned and faithfully reproduced the associations between signal segment positions and features of the meanings.

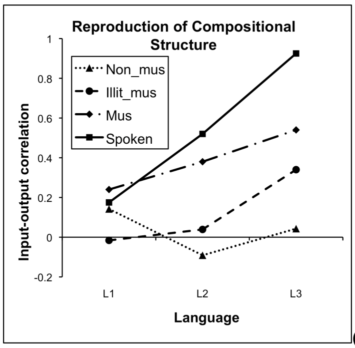
A two-way ANOVA of language (L1, L2, L3) and condition (non-musician, illiterate musician, musician, spoken) was carried out to assess the impact of condition and language on reproduction of signals and reproduction of structure. Results revealed significant effects of language ($p<0.001$ for signal reproduction; $p<0.01$ for structure reproduction), of condition ($p<0.001$ for both signal and structure reproduction) and a significant interaction for signal reproduction ($p<0.05$) but not for structure reproduction ($p=0.14$).

Further ANOVAs were then applied to pairs of conditions to obtain a more detailed picture. Comparing musicians with non-musicians reveals a strong effect of literacy and practice on the capacity to faithfully reproduce both the signals (effect of condition: $p=0.001$; effect of language: $p<0.01$; interaction $p<0.01$) and the compositional structure (effect of condition: $p<0.001$; non significant effect of language: $p=0.36$). This indicates that musicians were better

able than non-musicians to remember and produce signals and structure; performance overall improved as the degree of structure of the input languages increased; and, when reproducing the signals, musicians' advantage increases in the more structured languages. On the other hand, the compositional structure of the three languages was reproduced to similar degrees of accuracy.



(a)



(b)

Figure 2. Mean values of the reproduction of the signals and of the compositional structure of the input languages in the three languages and the four conditions. (a) Number of segments in participants' output languages that exactly matched the corresponding segment (in the same position) in their respective input languages. The theoretical maximum is 42, as we only take into account the 14 words participants were exposed to – remember a 50% bottleneck was in place during training. (b) Correlation (Pearson's r) of the nine partial *RegMaps* (which define the compositional structure of the language) in the input and output languages.

To separate the effects of practice from those of literacy, we compare the performance of illiterate musicians (who have practice but not literacy) with non-musicians (who have neither) and with musicians (who have both). Comparing illiterate musicians with musicians shows an effect of literacy both on reproduction of signals (effect of condition: $p<0.01$; effect of language: $p<0.01$; no significant interaction) and reproduction of compositional structure (effect of condition: $p<0.05$; effect of language: $p=0.05$; no significant interaction). Comparing illiterate musicians with non-musicians reveals an effect of practice on reproduction of signals (effect of condition: $p<0.05$; effect of language: $p=0.21$; no significant interaction) but not on reproduction of the compositional structure (effect of condition: $p=0.38$; effect of language: $p=0.18$). This indicates that practice and familiarity with music provides an advantage for reproducing the musical tunes, but not for noticing and reproducing the compositional structure of the system. For this task, literacy is required.

Comparing the results from the two language modalities (musical versus spoken) revealed an effect of modality of language on reproduction of signals

(effect of modality $p<0.05$; effect of level of language structure: $p<0.05$; no significant interaction), indicating that the spoken signals were much better learned and reproduced than the musical tunes, even in the absence of structure (see results for Language 1 in Fig. 2). This may be due to a mismatch between the conditions: while musicians are proficient in reading and playing music, they do not produce musical vocal output as frequently as the students produce spoken vocal output. As for the ability to reproduce the compositional structure of the language, very interestingly, modality had no impact (p=0.20), but the level of language structure did (p<0.01). This suggests that musicians learned the musical languages' structure as competently as undergraduates learned the spoken language's structure, and they did so despite the fact that while spoken signals usually map to referential meaning, musical signals usually do not. This result strongly supports out hypothesis that literacy promotes learning and reproduction of compositional structure.

2.6. *Conclusion*

Adding to the literature on the effects of literacy on language processing, our results show how literacy impacts on learning and reproduction not only of the signals, but of the *compositional structure* of a miniature artificial language. While all the participants in our experiment were able to name the objects in their artificial languages, only the signals produced by (musically or orthographically) literate learners retained the compositional relationship with the meanings that was present in the languages they were exposed to. In other words, only literates were able to stably replicate the structure of their input languages.

Language structure evolves as it is transmitted over the generations, and is affected by individual learning biases. The present results suggest that literacy and training facilitate learning and reproduction of compositional structure, supporting the hypothesis that these factors biased the evolution of linguistic structure in the direction of increased compositionality.

Acknowledgments

Monica Tamariz was funded by a Leverhulme Trust Postdoctoral Fellowship, ESRC Grant RES-062-23-1537 and AHRC Grant AH/F017677/1; J. Erin Brown was funded by Scottish Funding Council and the University of Edinburgh's College of Humanities and Social Science and Keelin M. Murray was funded by ESRC Postgraduate Studentship ES/H012605/1.

References

Cornish, H. Tamariz, M. and Kirby, S. (in press). 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*: 59:4S1.

de Villiers, P.A. & de Villiers, J.G. (1972). Early judgment of syntactic and semantic acceptability by children. *Journal of Psycholinguistic Research,* 1(4), 294-310.

Deacon, T. (1997). They symbolic species. New York: Norton & Co.

Ehri, L.C. (1985). Effects of printed language acquisition on speech. In D.R. Olson, N. Torrance & A. Hildyard (eds.), *Literacy, language, and learning: The nature and consequences of reading and writing*, pp. 333–367. Cambridge: Cambridge University Press.

Kalish, M., Griffiths, T. & Lewandowsky, S. (2007). Iterated learning: Intergenerational knowledge transmission reveals inductive biases. *Psychonomic Bulletin and Review*, 14(2), 288-294.

Karanth, P. & Suchitra, M.G., (1993). Literacy acquisition and grammaticality judgments in children. In: Scholes, R.J. (ed.), *Literacy and language analysis*, Hillsdale, NJ: Erlbaum.

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

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(31), 10681-10686.

Nunes, T., Bryant, P. & Bindman, M. (2006). The effects of learning to spell in children's awareness of morphology. *Reading and Writing*, 19, 767-787.

Olson, D.R. (1996). Towards a psychology of literacy: on the relations between speech and writing. *Cognition*, 60, 83-104.

Ramachandra, V. & Karanth, P. (2007). The role of literacy in the conceptualization of words: Data from Kannada-speaking children and non-literate adults. *Reading and Writing*, 2(3), 173-199.

Schoeneman, P.T. (1999). Syntax as an emergent characteristic of the evolution of semantic complexity. *Minds and Machines,* 9, 309-346.

Tamariz, M. (in press). Could arbitrary imitation and pattern completion have bootstrapped human linguistic communication? *Interaction Studies.*

Tamariz, M. & Smith, A.D.M. (2008). Quantifying the regularity of the mappings between signals and meanings. In A. D. M. Smith, K. Smith & R. Ferrer i Cancho (eds.): *The Evolution of Language: Proceedings of the 7th International Conference on the Evolution of Language*, pp. 315-322.