ASSESSING MEASURES OF MORPHOLOIGCAL COMPLEXITY AS PREDICTORS OF LEARNING BY NEURAL NETWORKS AND HUMANS

TAMAR JOHNSON[*1], KEXIN GAO[1], JENNIFER CULBERTSON[1], HUGH RABAGLIATI[2] and KENNY SMITH[1]

[*]Corresponding Author: tamarjohn@gmail.com
[1]Centre for Language Evolution, University of Edinburgh, UK
[2]Department of Psychology, University of Edinburgh, UK

Morphological paradigms differ widely across languages: some feature relatively few contrasts, and others, dozens. Under the assumption that languages evolve to maximise their learnability and that simpler systems are generally easier to learn (e.g. Chater & Vitanyi, 2003), this variation is surprising. Recent work on morphological complexity has resolved this paradox by arguing that certain features of even very large paradigms make them easy to learn and use. Specifically, Ackerman and Malouf (2013) propose an information-theoretic measure, i-complexity, which captures the extent to which forms in one part of a paradigm predict others. They contrast this measure with e-complexity, which is commonly used as a measure of morphological complexity in the literature (e.g., Bickel & Nichols, 2005) and captures the number of distinctions made by the language and the different ways to mark each grammatical function. They show that languages which differ widely in their e-complexity exhibit similarly low i-complexity; this suggests that having predictive relationships between inflections (i.e. low i-complexity) reduces the learnability challenge for learners even when the morphological paradigm makes many contrasts (i.e. has high e-complexity). Here, we test these measures of morphological complexity across two experiments, to evaluate whether i-complexity in fact influences the learnability of morphological paradigms as predicted by Ackerman & Malouf.

In Experiment 1 we tested, with recurrent neural networks (RNNs) and human participants trained on artificial languages, whether i-complexity predicts the learnability of inflectional paradigm. Learners were trained on an artificial language consisting of nine nouns divided into three noun classes and marked for three numbers: singular, dual and plural (Table 1). We created two languages, matched on e-complexity but differing in i-complexity. Only in the low i-

complexity paradigm, the singular form of a word always predicts its form in dual. Learning in this experiment was staged - learners were first exposed to singular forms, then plural, then dual; staging allows us to test whether having learned the singular facilitates the later learning of the dual in the paradigm where the singular is predictive. Results show that for the RNNs, i-complexity affected the speed of learning, with the low i-complexity paradigm being learnt more rapidly; however, for human participants there was only weak evidence for this effect.

Table 1. Example paradigm for low i-complexity (left) and high i-complexity languages (right).

	Singular	*Dual*	*Plural*
1	-op	-um	-ib
2	-at	-oc	-el
3	-op	-um	-od

	Singular	*Dual*	*Plural*
1	-op	-um	-ib
2	-at	-um	-el
3	-op	-oc	-od

In Experiment 2 we manipulated both i-complexity and e-complexity to test whether 1) the effect of i-complexity found in Experiment 1 holds when learning is not staged (i.e., learners are exposed to wordforms from the entire paradigm randomly), and 2) whether e-complexity has an effect on learning. For RNNs there was no significant difference in learning the low vs. high *i-complexity* paradigms. However, there was a significant difference in learning the low vs. high *e-complexity* paradigms: neural networks trained on the low e-complexity paradigms achieved higher accuracy. This was mirrored in human participants: there was no effect of i-complexity but a significant effect of e-complexity. These results suggest that the benefits of lower i-complexity are dependent on learners' prior knowledge of predictive forms (not necessarily the case in natural language learning). By contrast lower values of e-complexity are advantageous for learning independent of when different parts of the paradigm are learned.

Taken together, these results cast doubt on Ackerman and Malouf's hypothesis that i-complexity rather than e-complexity drives the learnability of morphological paradigms. Since i-complexity and e-complexity are inversely correlated, the limits on i-complexity they identify in natural languages likely reflect how the two measures relate rather than how i-complexity shapes the evolution of morphological paradigms.

References

Ackerman, F., & Malouf, R. (2013). Morphological organization: The low conditional entropy conjecture. *Language*, 429-464.

Bickel, B., & Nichols, J. (2005). Inflectional synthesis of the verb. *The world atlas of language structures*, 94-97.

Chater, N., & Vitányi, P. (2003). Simplicity: A unifying principle in cognitive science?. *Trends in cognitive sciences*, 7(1), 19-22.