The length of words reflects their conceptual complexity

Molly Lewis and Michael C. Frank

Psychology Department, Stanford University,
450 Serra Mall, Stanford, CA 94305, USA

*To whom correspondence should be addressed; E-mail: mll@stanford.edu.

Are the forms of words systematically related to their meaning? The arbitrariness of the sign has long been a foundational part of our understanding of human language (*1, 2*). Theories of communication predict a relationship between form and meaning, however: longer descriptions should convey more complex meanings (*3, 4*). Here we show that the lexicons of human languages reflect this relationship between linguistic and cognitive complexity. We asked participants to rate the conceptual complexity of word meanings and found that their judgements correlated highly with word length across XX languages, even controlling for frequency and concreteness. This relationship is productively encoded in the minds of speakers, as well: Adults and children both mapped longer words to more complex meanings, and more complex meanings to longer words, in comprehension and production tasks and across a wide range of stimuli. In addition, explicit judgments of complexity were highly correlated with an implicit measure of study time in a memory task, suggesting that complexity is directly related to basic cognitive processes. These results

point to a general regularity in the design of lexicons and suggests the importance of cognitive constraints on language evolution (*5, 6*).

Human languages universally contain sequences of sounds — words — that are associated with particular meanings. A foundational part of our understanding of human language is that these associations are arbitrary (*?, 1*). This assumption is supported by a superficial survey of word forms across languages: different languages use different words to refer to similar meanings. However, several theories of communication predict a systematicity in these mappings (*3, 4*). They predict that longer utterances should be associated with more complex meanings (a complexity bias). Here we address whether this bias is present in language at the level of words. We first examine whether speakers have a lexical complexity bias that is productive by asking how speakers interpret novel words. Given evidence for a productive complexity bias in words, we then ask whether this systematicity is encoded in the lexicon of XX languages.

In Study 1, we asked whether speakers show a bias to map a relatively long novel word onto a relatively more complex referent. Participants were presented with

From an information theoretic perspective,

In Study 1, we asked whether speakers norms + mapping

In Study 2, norms + mapping + production (a single sentence)

In Study , we did complexity norms. Highly correlated.

In Study 4, cross linguistic.

Figs. (1) six panels with sorted display of images on left (geons + objects) and effect size plots on right (complexity + RT) (2) bar graph of correlations (study 4) (with parial correlations, mono morphemic, and open class only)

Tables

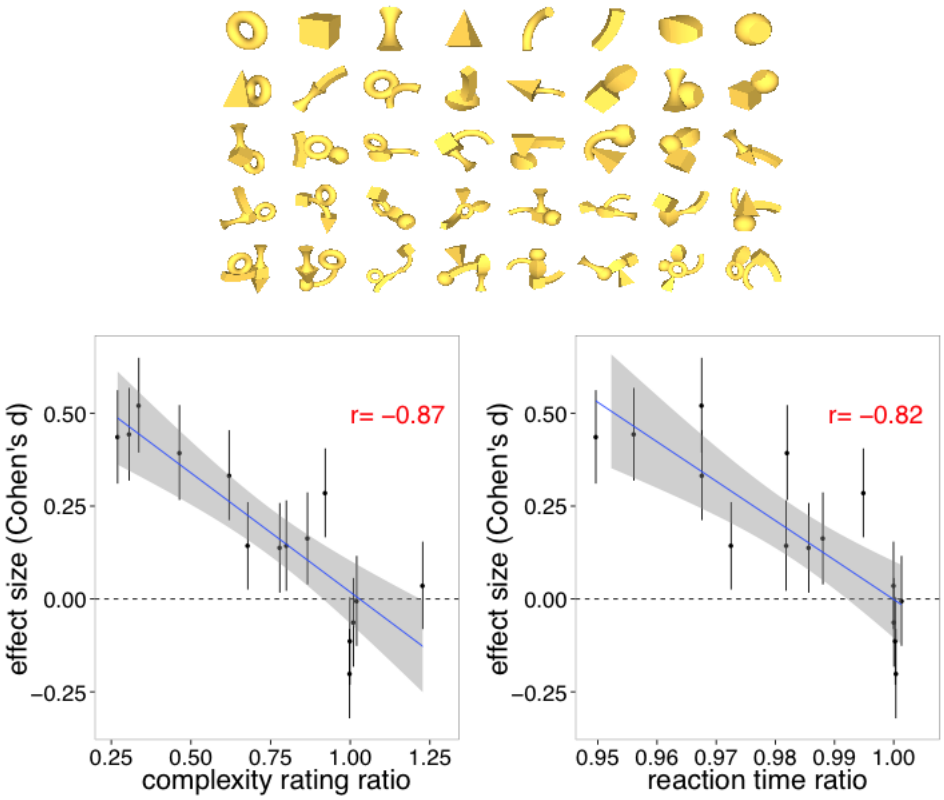


Figure Legends

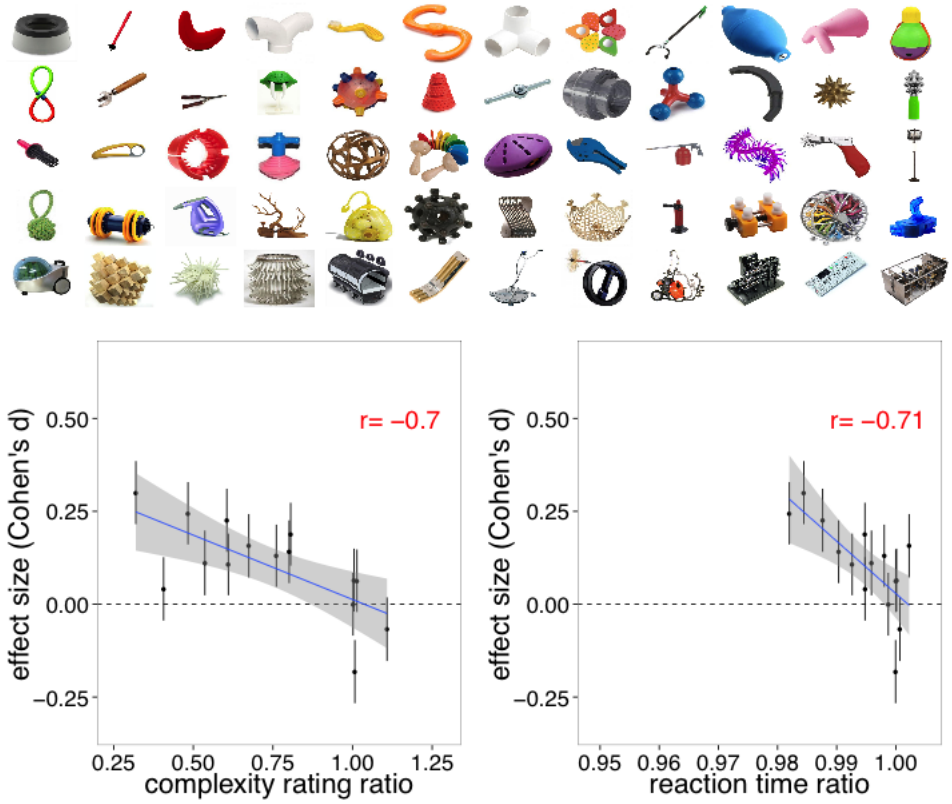
Methods

References

Acknowledgements

Author Contributions

Author Information



References and Notes

1. F. Saussure, *London: Peter Owen* (1916).

2. C. Hockett, *Scientific American* **203**, 88 (1960).

3. L. Horn, *Meaning, form, and use in context* **42** (1984).

4. T. Jaeger, R. P. Levy, *Advances in neural information processing systems* (2006), pp. 849–856.

5. M. H. Christiansen, N. Chater, *Behavioral and brain sciences* **31**, 489 (2008).

6. E. Lieberman, J.-B. Michel, J. Jackson, T. Tang, M. A. Nowak, *Nature* **449**, 713 (2007).