THE GENERALITY OF FORM-TO-MEANING ICONICITY

CHARLES P. DAVIS[*1], HANNAH M. MORROW[1], MARGHERITA DE LUCA[2], and GARY LUPYAN[3]

[*]Corresponding Author: charles.davis@uconn.edu
[1]Department of Psychological Sciences, University of Connecticut, Storrs, USA
[2]Department of Document Studies, Linguistics, Philology and Geography, Sapienza University of Rome, Rome, Italy
[3]Department of Psychology, University of Wisconsin-Madison, Madison, USA

Do certain words fit some meanings better than others? If so, to what extent do people make use of these relationships? Given a choice of whether a "nonsense" word refers to one or another object, people show predictable choices (e.g., Sapir, 1929; see Dingemanse et al., 2015 for review). Understanding the generality and origins of iconicity – a resemblance between a word's form and its meaning – is critical for research in language evolution as it lends insight into how initial form-to-meaning links may be established prior to emergence of large-scale conventional vocabularies.

To test whether form–meaning resemblances affect behavior in a relatively open-ended task, in a recent study (Davis et al., 2019), we first asked people to match written English-like nonsense words (from Westbury et al., 2018) to properties, e.g. is a *horgous* large, round, etc. (Fig. 1). A second group drew creatures in response to the same nonce words. A third group was asked to indicate whether the drawn creatures were large, round, etc. A final group was shown the creatures and asked to match them to nonsense words. Remarkably, the form of the nonsense words permeated the creature drawings: people judged the drawn creatures as having the same properties connoted by the original nonce words and were able to match the drawn creatures back to the name used to elicit them with higher-than-chance accuracies (e.g. matching the *horgous* drawings back to the word "horgous").

What explains these associations? One possibility is that they are mediated by idiosyncratic similarity to real English words, e.g., a "horgous" may be *large* due to form overlap with "humongous." Alternatively, the form–meaning associations may reflect more general sources of information, e.g., experience associating perceptual inputs across modalities (Lupyan & Casasanto, 2015).

To distinguish between these possibilities, we recruited native Spanish and Italian speakers to match the nonce words to translated properties. Fig. 1 shows how similar the results were across languages (cross-language correlations: r = .75–.91). Some discrepancies (e.g., *cougzer* as feminine vs. masculine) likely stem from overlap with real-word neighbors, but these do not explain the overwhelming similarity, as the orthographic neighborhoods of the nonce words in each language differ considerably.

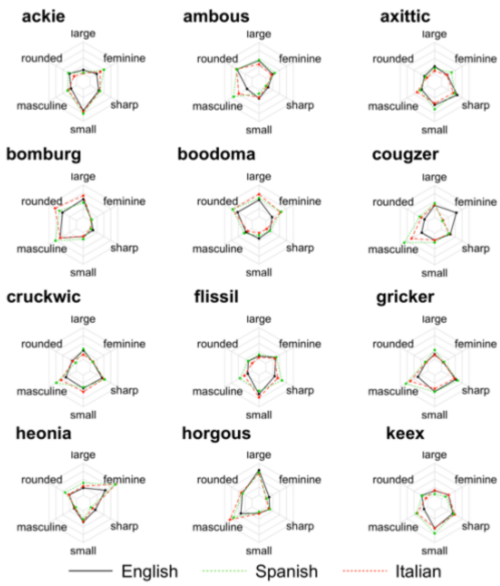

Fig. 1. Nonce-word property ratings for three languages.

We next examined whether the form–meaning associations are present in the distributed structure of each language. We trained word embedding models using the fast-text algorithm (Bojanowski et al., 2016) trained on English, Spanish, and Italian Wikipedia as well as parallel-translations of the OpenSubtitles corpus. We used the nonce words as input to the model and examined the proximity of resulting semantic representations to each property (e.g., *large*). The models were correlated with human ratings (z > 10). Strikingly, predictions were stronger for *between*-language pairs (e.g., English responses were more poorly predicted by English embeddings than by Spanish and Italian embeddings). This is unexpected if iconic associations derive from nearby real words, as such proximity-driven responding would increase within-language matches.

Nonarbitrary form–meaning associations appear to be surprisingly potent, influencing even open-ended drawing tasks. When matching nonce words to properties, English, Italian, and Spanish speakers show strikingly similar behavior (though replication to more diverse languages is clearly needed, as the languages here were chosen for convenience). Lastly, behavior was predicted by large-scale form–meaning associations in language as learned by a simple neural network.

References

Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2017). Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, *5*, 135–146.

Davis, C. P., Morrow, H. M., & Lupyan, G. (2019). What does a horgous look like? Nonsense words elicit meaningful drawings. *Cognitive Science*, *43*(10), e12791.

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. *Trends in Cognitive Sciences*, *19*(10), 603–615.

Lupyan, G., & Casasanto, D. (2015). Meaningless words promote meaningful categorization. *Language and Cognition*, *7*(2), 167–193.

Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology*, *12*, 225–239.

Westbury, C., Hollis, G., Sidhu, D. M., & Pexman, P. M. (2018). Weighing up the evidence for sound symbolism: Distributional properties predict cue strength. *Journal of Memory and Language*, *99*, 122–150.