Sound-meaning systematicity in early word learning

Justin B. Kueser and Laurence B. Leonard (Purdue University)
jkueser@purdue.edu

Early word learning relates both to a word's semantic content and its phonological form. Young children tend to learn highly salient, concrete words made up of sounds within their phonological inventories before other words. Children also attend to how semantics and phonology systematically vary across words and use this information as a cue to support word learning in controlled, experimental settings across a wide range of ages (e.g., Imai, Kita, Nagumo, & Okada, 2008). Research has shown that sound-meaning systematicity and a related concept, iconicity, are related to naturalistic word learning, with more systematic or iconic words demonstrating younger ages of acquisition (e.g., Monaghan, Shillcock, Christiansen, & Kirby, 2014; Thompson, Vinson, Woll, & Vigliocco, 2012). These studies have treated systematicity and iconicity as a word-level value that does not change based on an individual child's vocabulary; that is, in these studies, each word possesses a fixed value for systematicity or iconicity. Nevertheless, we know that the ability to take advantage of sound symbolism varies as a function of age (e.g., in learning novel verbs conforming to the Japanese sound symbolism system as shown in Imai et al., 2008), suggesting that a child's individual vocabulary size and the particular words known will affect the degree to which the child responds to sound-meaning systematicity. Each word's contribution to overall sound-meaning systematicity within an individual child will change based on the other words the child knows.

We sought to address this potential problem by investigating the degree of sound-meaning systematicity in individual children's vocabularies. Working from analyses suggesting that the English language as a whole demonstrates more sound-meaning systematicity than would be expected by chance, we investigated the role that sound-meaning systematicity plays in word learning in a large sample of administrations of MacArthur-Bates Communicative Development Inventories in American English, British English, and Mexican Spanish. Phonological and semantic similarity values were calculated for all pairs of words in each language and these values were used to calculate sound-meaning systematicity for each child's individual vocabulary. We found that in each of the languages, across a wide range of vocabulary sizes, children's vocabularies tended to demonstrate more sound-meaning systematicity than would be expected based on word frequency, word length, part of speech, phonotactic probability, neighborhood density, and consonant age-of-acquisition (see Figure 1). Further, we found a significant degree of individual variation in sound-meaning systematicity across children, especially at smaller vocabulary sizes. We also found that sound-meaning systematicity tended to be present at similar levels in each of the languages at smaller vocabulary sizes but tended to diverge in different languages at larger vocabulary sizes. These cross-linguistic patterns appeared to follow the overall trends for sound-meaning systematicity in the full adult vocabulary of each language that were observed in Dautriche, Mahowald, Gibson, and Piantadosi (2017), suggesting that from early development, children exploit the useful information provided by sound-meaning systematicity to a degree that reflects its overall presence in the adult language. The results also suggest that the use of sound-meaning systematicity may be an important factor to investigate in future studies of children's vocabulary growth.

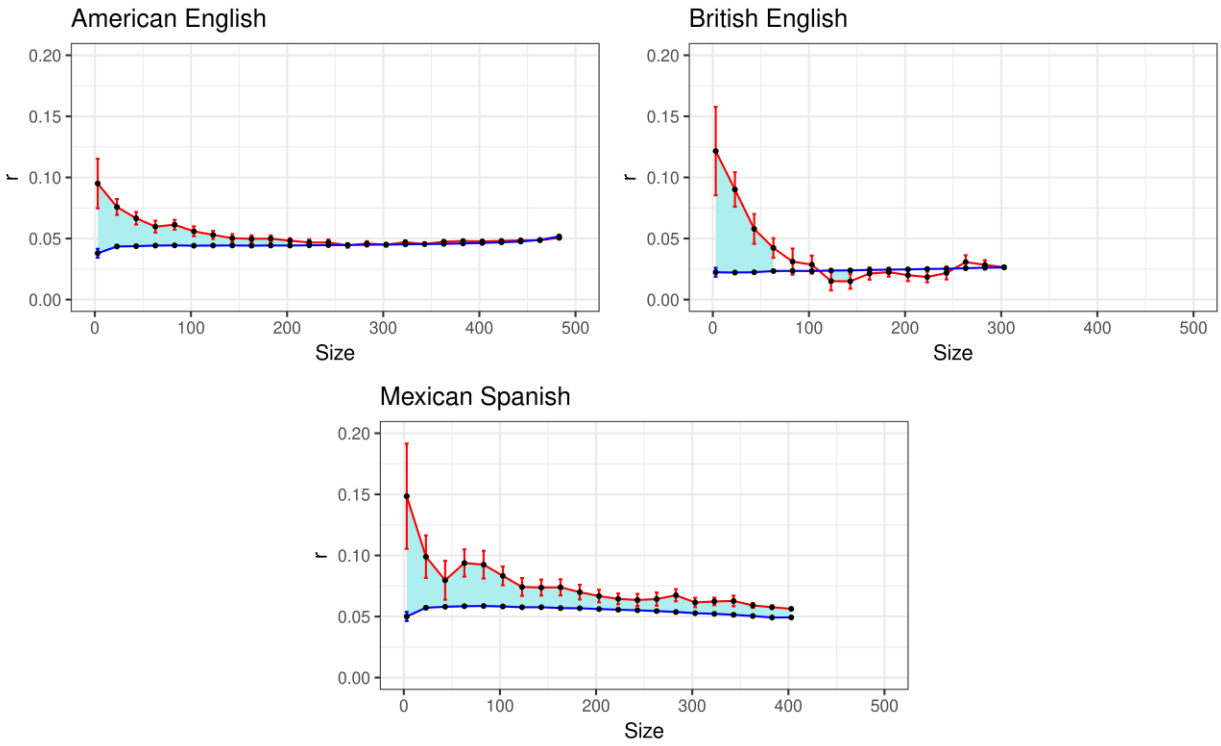

Figure 1. Sound-meaning correlations grouped in bins of 20 words. Dark blue lines indicate values for random weighted vocabularies, while red lines indicate values for children's observed vocabularies. The children's systematicity values were compared to the systematicity values of randomly generated vocabularies that were weighted to control for word-level properties (see abstract text). Error bars represent 95% confidence intervals. Light blue areas represent clusters of significantly different ($p < .05$) values after applying a bootstrapped cluster-based permutation test.

References

Dautriche, I., Mahowald, K., Gibson, E., & Piantadosi, S. T. (2017). Wordform similarity increases with semantic similarity: An Analysis of 100 Languages. *Cognitive Science, 41*(8), 2149–2169. doi:10.1111/cogs.12453

Imai, M., Kita, S., Nagumo, M., & Okada, H. (2008). Sound symbolism facilitates early verb learning. *Cognition, 109*(1), 54–65. doi:10.1016/j.cognition.2008.07.015

Monaghan, P., Shillcock, R. C., Christiansen, M. H., & Kirby, S. (2014). How arbitrary is language? *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1651), 20130299–20130299. doi:10.1098/rstb.2013.0299

Thompson, R. L., Vinson, D. P., Woll, B., & Vigliocco, G. (2012). The road to language learning is iconic: Evidence from British Sign Language. *Psychological Science, 23*(12), 1443–1448. doi:10.1177/0956797612459763