

## Comprehension of Non-Linguistic Vocalizations across Cultures

Aleksandra Ćwiek<sup>1</sup>, Susanne Fuchs<sup>1</sup>, Cornelia Ebert<sup>1,2</sup>, Manfred Krifka<sup>1</sup>,  
Christoph Draxler<sup>3</sup>, Bodo Winter<sup>4</sup> and Marcus Perlman<sup>4</sup>

<sup>1</sup> Leibniz-Zentrum Allgemeine Sprachwissenschaft, Berlin, <sup>2</sup> Goethe-Universität Frankfurt,  
<sup>3</sup> IPS, Ludwig-Maximilian-Universität, München, <sup>4</sup> University of Birmingham

It is widely observed that visual communication with gestures enables successful communication between speakers of different languages. When people lack common words to convey their meaning, they are able to improvise iconic gestures to ground communication about varied domains like actions, shape and size, and spatial relationships. This remarkable human ability is the foundation of the parlor game ‘charades,’ and it is thought to play an important role in language acquisition and evolution. But can people also innovate iconic vocalizations to communicate meanings? To find out, Perlman and Lupyan (2018) conducted a contest with a \$1000 prize that invited participants to create novel vocalizations—no words permitted—for 30 different meanings. The meanings spanned actions, humans, animals, inanimate objects, properties, quantifiers, and demonstratives. The vocalizations were evaluated by the ability of naive listeners to guess their intended meanings from multiple alternatives, with the winning submission determined as the one that was guessed most accurately. Overall, guessing accuracy was well above chance for most of the submissions and across almost all of the meanings. Thus, the results suggested that people are able to communicate successfully with iconic vocalizations about various concepts, without the use of words. However, strong evidence for this claim depends on the potential for vocal communication across widely disparate cultures and linguistic backgrounds. Perlman and Lupyan’s contestants were English speakers, and their listeners were all American English speakers, raising the possibility that culture-specific cues were used by listeners in determining the intended meanings.

The current project is investigating whether the vocalizations produced for the contest can be understood by people from a wide range of cultural and linguistic backgrounds. We translated and conducted comprehension surveys with listeners from 28 diverse languages, and from 10 language families using Percy (Draxler, 2011), a web-tool for experiments. Twenty-four surveys yielded enough participants. For each meaning, the participants listened to the three vocalizations that were guessed most accurately in the English experiment, and for each, guessed its meaning from among six alternatives. Given that the vocalizations were produced by English speakers, we predicted that the response accuracy of listeners would depend on the genetic distance of the given language from English, which we took as a rough approximation of cultural and linguistic overlap. Here, we simplify it to a binary distinction between Indo-European and Non-Indo-European languages. The preliminary results of a Bayesian mixed logistic regression model show that speakers of all examined languages perform significantly above chance level (16.7%). The accuracies range from the lowest of Thai, with 51%, to the highest of German, with 72%. The data show that the linguistic genetic distance variable is not enough to predict the guessing accuracy. In fact, listeners from the Ugro-Finnic language family and Korean listeners performed better than other languages more closely related to English. We think that the education level or exposure to western culture may pose a missing link (see Heinrich et al., 2010). As for the meanings we tested, overall, the guessing accuracy was higher than chance for all 30 meanings. Guessers across all languages were most successful with actions and animate nouns, which corroborates previous results (e.g., Winter et al., 2017; Perlman et al., 2018).

To enable testing of remote populations, we also designed a simplified version where orthography was not used. Therefore, only nouns were tested as they are the most depictable without writing. For each of the 12 nouns, we used a picture and placed it in front of the participant on the table. Same as in the web version, three vocalizations were played for each noun in a pre-randomized order. In this version of the experiment, we tested a group from Vanuatu (Daakie), two remote groups from Brazil (Brazilian Portuguese and Palikur), a group from Morocco (Tashlhiyt Berber), and as a comparison a group of German speakers and two English speaking groups (UK and USA). On average, responses across the seven languages were 55% correct, much above chance level (8%). All languages performed above chance and the accuracy ranged from the lowest of 34% for Portuguese to the highest of German (63%). Across the board, all meanings were guessed above chance level, ranging from the lowest of 27.5% (“fruit”) to the highest of 86% (“child”). On average, responses to animate items (72%) were much more correct than responses to inanimate items (39%).

Our results provide the first strong evidence that people are able to comprehend improvised iconic vocalizations for successful, cross-linguistic communication about a variety of

meanings. They also point to semantic domains where iconic vocalizations may be less effective. Altogether, they speak to the universal potential for vocal iconicity to scaffold communication between speakers without a common language and thus ground the formation of spoken symbols.

## References

- Draxler, C. (2011). Percy-An HTML5 framework for media rich web experiments on mobile devices. In *Twelfth Annual Conference of the International Speech Communication Association*.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2-3), 61-83.
- Perlman, M., Little, H., Thompson, B., & Thompson, R. L. (2018). Iconicity in Signed and Spoken Vocabulary: A Comparison Between American Sign Language, British Sign Language, English, and Spanish. *Frontiers in Psychology*, *9*, 1433.
- Perlman, M., & Lupyan, G. (2018). People can create iconic vocalizations to communicate various meanings to naïve listeners. *Scientific Reports*, *8*(1), 2634.
- Winter, B., Perlman, M., Perry, L. K., & Lupyan, G. (2017). Which words are most iconic? *Interaction Studies*, *18*(3), 443-464.