A ‘size meets cuteness’ relation in German vowels

Dominic Schmitz
Heinrich-Heine-Universität Düsseldorf

Sound symbolism is a specific form of cross-modal correspondence: Certain sounds become meaningful when they are combined with other sensory information. Prominent types of such sensory information are the rather easily measurable size and the more complex concept of cuteness. However, to date, no combined account of size and cuteness and potential interactions thereof has been proposed.

Size has been under investigation in multiple studies on sound symbolism during the last decades, e.g. [1]–[4]. Across studies, it was found that open vowels correlate with bigger size, while closed vowels correlate with smaller size. Ohala [5] assumed this finding to be derived from the fact that smaller things typically produce higher-frequency sounds, leading to vowels with pertinent characteristics to be correlated with smallness.

Cuteness as a concept can be understood as a more complex form of simple geometric shape, as was the feature of focus in previous studies, e.g. [4], [6], [7]. Cuteness, especially from its biological perspective as comprised in the so-called ‘baby schema’ [8], is a fundamental feature of human perception and correlates, among other things, with size [9]. Research on Japanese has shown that cuteness is also found as sensory information to be combined with speech sound [10].

Bringing size and cuteness together, the present paper aims to establish a multidimensional relation from small to big and from not cute to cute for long vowels of Standard German (i.e. /aː, ɛː, eː, iː, oː, øː, uː, yː/), providing further insight into the nature of sound symbolism.

For this, a forced-choice task was conducted using OpenSesame [11] online. As auditory stimuli, disyllabic pseudowords were used to control for potentially confounding lexical [12], [13] and contextual [14], [15] effects. For both syllables of a pseudoword, the nucleus consisted of one of the vowels under investigation, while the onsets of the syllables consisted of 12 combinations of /d, f, j, k/ and /r/. In total, 96 pseudowords, i.e. 12 per vowel, were used. As visual stimuli, images of phantasy creatures [16] were used. For each trial, participants were shown five differently sized versions of a randomly chosen creature. The task was to decide which image version matched the audio stimulus of a trial best. As cuteness judgements most likely differ by participants, after the forced-choice part of the experiment, participants were again shown all creature images to judge them for their cuteness on a five point scale.

The size response then entered a generalised additive mixed model regression analysis as dependent variable. Cuteness judgments, vowel quality, as well as several control variables (e.g. coda consonants, phonological neighbourhood density) were introduced as independent variables, while participant ID was included as random effect. Overall, /aː/ is considered bigger than all other vowels, while cuteness judgement ratings do not show a significant effect on their own. However, having vowel quality and cuteness judgements interact, a noteworthy pattern emerges: For the open vowel /aː/ and for the close vowels /uː, iː/, the interaction reaches significance. While the size judgements for /aː/ further increase with cuteness, the size judgements for /uː, iː/ further decrease.

The 'size meets cuteness’ relation found in the present paper offers an insight to how structures of sound symbolism interact with each other. Sound symbolism, while seemingly simple when considering but one type of sensory information, manifests as an intricate interaction when different types of sensory information are available. The present findings contribute to the growing body of evidence for and the nature of sound symbolism and call for the incorporation of multiple sources of sensory information where applicable.

doi:10.11576/pundp2022-1039

[1] B. Berlin, "The First Congress of Ethnozoological Nomenclature," *Journal of the Royal Anthropological Institute*, vol. 12, no. s1, pp. S23–S44, Mar. 2006, doi: 10.1111/j.1467-9655.2006.00271.x.
[2] D. E. Blasi, S. Wichmann, H. Hammarström, P. F. Stadler, and M. H. Christiansen, "Sound-meaning association biases evidenced across thousands of languages," *Proc Natl Acad Sci U S A*, vol. 113, no. 39, pp. 10818–10823, Sep. 2016, doi: 10.1073/PNAS.1605782113/SUPPL_FILE/PNAS.201605782SI.PDF.
[3] N. Johansson, "Tracking linguistic primitives," Sep. 2017, doi: 10.1075/ILL.15.03JOH.
[4] C. Westbury, G. Hollis, D. M. Sidhu, and P. M. Pexman, "Weighing up the evidence for sound symbolism: Distributional properties predict cue strength," *Journal of Memory and Language*, vol. 99, pp. 122–150, Apr. 2018, doi: 10.1016/J.JML.2017.09.006.
[5] J. J. Ohala, "The frequency code underlies the sound-symbolic use of voice pitch," *Sound Symbolism*, pp. 325–347, Aug. 1995, doi: 10.1017/CBO9780511751806.022.
[6] A. J. Bremner, S. Caparos, J. Davidoff, J. de Fockert, K. J. Linnell, and C. Spence, "'Bouba' and 'Kiki' in Namibia? A remote culture make similar shape–sound matches, but different shape–taste matches to Westerners," *Cognition*, vol. 126, no. 2, pp. 165–172, Feb. 2013, doi: 10.1016/J.COGNITION.2012.09.007.
[7] A. D'Onofrio, "Phonetic Detail and Dimensionality in Sound-shape Correspondences: Refining the Bouba-Kiki Paradigm:," *http://dx.doi.org/10.1177/0023830913507694*, vol. 57, no. 3, pp. 367–393, Nov. 2013, doi: 10.1177/0023830913507694.
[8] V. Lehmann, E. M. J. Huis in't Veld, and A. J. J. M. Vingerhoets, "The human and animal baby schema effect: Correlates of individual differences," *Behavioural Processes*, vol. 94, pp. 99–108, Mar. 2013, doi: 10.1016/J.BEPROC.2013.01.001.
[9] M. L. Kringelbach, E. A. Stark, C. Alexander, M. H. Bornstein, and A. Stein, "On Cuteness: Unlocking the Parental Brain and Beyond," *Trends in Cognitive Sciences*, vol. 20, no. 7, pp. 545–558, Jul. 2016, doi: 10.1016/j.tics.2016.05.003.
[10] G. Kumagai, "A sound-symbolic alternation to express cuteness and the orthographic Lyman's Law in Japanese," *Journal of Japanese Linguistics*, vol. 35, no. 1, pp. 39–74, May 2019, doi: 10.1515/JJL-2019-2004/MACHINEREADABLECITATION/RIS.
[11] S. Mathôt, D. Schreij, and J. Theeuwes, "OpenSesame: An open-source, graphical experiment builder for the social sciences," *Behavior Research Methods*, vol. 44, no. 2, pp. 314–324, Jun. 2012, doi: 10.3758/s13428-011-0168-7.
[12] N. K. Caselli, M. K. Caselli, and A. M. Cohen-Goldberg, "Inflected words in production: Evidence for a morphologically rich lexicon," *Quarterly Journal of Experimental Psychology*, vol. 69, no. 3, pp. 432–454, 2016, doi: 10.1080/17470218.2015.1054847.
[13] S. Gahl, "Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech," *Language (Baltim)*, vol. 84, no. 3, pp. 474–496, 2008, doi: 10.1353/lan.0.0035.
[14] D. H. Klatt, "Linguistic uses of segmental duration in English: Acoustic and perceptual evidence," *J Acoust Soc Am*, vol. 59, no. 5, p. 1208, 1976, doi: 10.1121/1.380986.
[15] C. W. Wightman, S. Shattuck-Hufnagel, M. Ostendorf, and P. J. Price, "Segmental durations in the cicinity of prosodic phrase boundaries," *Journal of the Acoustical Society of America*, vol. 91, no. 3, pp. 1707–1717, 1992, doi: 10.1121/1.402450.
[16] R. van de Vijver and D. Baer-Henney, "Developing biases," *Frontiers in Psychology*, vol. 5, Jun. 2014, doi: 10.3389/fpsyg.2014.00634.

doi:10.11576/pundp2022-1039