



<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper published in . This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the original published paper (version of record):

Ekström, A G. (2022)

What's next for size-sound symbolism?

Frontiers in Language Sciences, 1

<https://doi.org/10.3389/flang.2022.1046637>

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-323658>



OPEN ACCESS

EDITED BY

Moreno I. Coco,
Sapienza University of Rome, Italy

REVIEWED BY

Elliot Murphy,
University College London,
United Kingdom

*CORRESPONDENCE

Axel G. Ekström
axeleks@kth.se

SPECIALTY SECTION

This article was submitted to
Psycholinguistics,
a section of the journal
Frontiers in Language Sciences

RECEIVED 16 September 2022

ACCEPTED 18 October 2022

PUBLISHED 03 November 2022

CITATION

Ekström AG (2022) What's next for
size-sound symbolism?
Front. Lang. Sci. 1:1046637.
doi: 10.3389/flang.2022.1046637

COPYRIGHT

© 2022 Ekström. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

What's next for size-sound symbolism?

Axel G. Ekström*

Speech, Music and Hearing, KTH Royal Institute of Technology, Stockholm, Sweden

This text reviews recent research in phonetic size-sound symbolism – non-arbitrary attributions of size properties to speech acoustic properties. Evidence from a wide range of research works is surveyed, and recent findings from research on the relationships between fundamental frequency, vowel articulation, consonant articulation, phonation type, mora count, and phonemic position, are discussed. It is argued that a satisfactory explanatory model of phonetic size-sound symbolism should meet two criteria: they should be able to explain both (1) the relationship between size and speech acoustics (Association criterion), and (2) the inconsistent findings observed across languages in the relevant literature (the Inconsistency criterion). Five theories are briefly discussed: The frequency code, Embodied cognition, Sound-meaning bootstrapping, Sapir-Whorf hypotheses, and Stochastic drift. It is contended that no currently available explanatory model of size-sound symbolism adequately meets both criteria (1) and (2), but that a combination of perspectives may provide much of the necessary depth. Future directions are also discussed.

KEYWORDS

sound symbolism, speech perception, frequency code, psycholinguistics, auditory cognition, biology of speech, bioacoustics

Introduction

In phonetics, size-sound symbolism – the non-arbitrary association of speech acoustics relationships to estimates of speaker physical size – has been the subject of almost a century of research efforts (Sapir, 1929; Jespersen, 1933; Peterfalvi, 1965; Ultan, 1978; Ohala, 1984, 1994; Diffloth, 1994; Tsur, 2006). The present brief review centers on recent research findings on size-sound symbolism in experimental settings. The text summarizes findings from various aspects of speech acoustics, including fundamental frequency, and vowel and consonant articulation¹. In a later section, explanatory models with bearing on the phenomenon are evaluated. Future directions are discussed in the final section.

¹ The text as such is concerned with psychological attributions based on acoustic properties of speech sounds. Topics more typical of traditional linguistics (e.g., processing of sentences, syntax, etc.), are not included (see e.g., Altmann, 1998), nor are they of immediate interest to the topic of phonetic symbolism more generally.

Cues to size in frequencies, words, and elsewhere

Symbolism in fundamental frequency

Various researchers have investigated whether fundamental frequency of phonation (f_0) is cognitively associated with size (Ohala, 1984; McComb, 1991; Masataka, 1994). f_0 reflects the rate of vocal fold oscillation, corresponding to pitch in perception. Because longer vocal folds oscillate more slowly, vocal fold length exhibits a negative correlation with f_0 . Thus, on average, larger species do indeed tend to produce lower-frequency f_0 (Hauser, 1993), though the variable size of various laryngeal properties likely plays a greater role than body size per se (Titze, 1994; Garcia et al., 2017; Grawunder et al., 2018). However, *within* species, various studies, including research on both humans (Gunter and Manning, 1982; Wermke and Robb, 2010) and nonhuman animals (e.g., Masataka, 1994) have failed to observe any such relationship.

Because the larynx is sensitive to sex hormones (Newman et al., 2000), f_0 represents a sexually selected trait in humans (Delgado, 2006; Puts et al., 2016). Thus, when size- f_0 correlations are observed, relationships are often conditional, and typically more prominent in male speakers (see overviews by Ey et al., 2007; Riede and Brown, 2013). As testosterone thickens the laryngeal folds, taller and more high-testosterone men typically exhibit lower-frequency f_0 (González, 2004; Bruckert et al., 2006). Considered overall, however, the contribution of f_0 to size estimates is limited, with one meta-analysis finding that only 2% of men's height and 0.5% of women's height were explained by f_0 (Pisanski et al., 2014), though f_0 may still convey semantic information. For example, pitch has been shown to correspond perceptually to spatial elevation in research on music perception (Shintel et al., 2006; Küssner et al., 2014), prosody (Ekström et al., 2022), and speech more generally (Shinohara et al., 2020). Due to the spurious and complicated nature of the evidence, however, much work has instead concentrated on other aspects of speech signals as suggestive of size.

Symbolism in vowels

Vowel articulation is reliably modeled as series of alternately compressed tubes through which pulmonary air flow is expelled, resulting in variable vowel quality (Fant, 1960). For example, [a] is reducible to an initial compressed tube (where compression corresponds to constrictions on vocal tract air flow, e.g., by tongue position or lip rounding), and a subsequent open tube, resulting in a “back” vowel; for the “frontal” [i], the relationship between tubes is reversed. The acoustic results are formants – high-energy peaks in the frequency spectrum, corresponding to resonances in the vocal tract. Resulting vowel quality has been argued as suggestive of physical size, such that higher vowels

and front vowels are suggestive of smaller physical size, and back vowels and lower vowels are suggestive of larger size (Tarte, 1974; Ultan, 1978; Fitch, 1997; Owren et al., 1997; Knoeferle et al., 2017).

Empirically, back vowels have indeed been shown to be positively associated with size in German and Hungarian (Elsen et al., 2021), and in Chinese, English, Japanese, and Korean (Shinohara and Kawahara, 2010). Researching ethnozoological nomenclature for birds and fish across four Central and South American languages, Berlin (1994) found that names exhibited significant size-sound symbolism, such that names for smaller creatures contained disproportionate numbers of high-frequency vowels, i.e., [i], and names for larger creatures contained disproportionate numbers of low-frequency vowel sounds, i.e., [a] and [u]. Similarly, a recent lexicographic approach by Winter and Perlman (2021) found that sound structure was indeed indicative of semantic size in English size adjectives (but not for general words), in particular for vowels [i] (small), [I] (small), and [a] (large).

Notably, however, while effects of linguistic background are sometimes negligible (Hoshi et al., 2019; Elsen et al., 2021), investigations often present puzzling contradictory results, challenging any supposed universality. In an important potential counterexample, Difloth (1994) examined iconicity in the Bahnar language of Vietnam and found iconic values opposite those purported universal (Ohala, 1984): high vowels corresponded to larger expressives, and low vowels to smaller ones. Observations running counter to typical findings are rare, however (Lockwood and Dingemanse, 2015). More common is an apparent lack of findings when some experimental or observational methods is applied to multiple languages. For example, while Berlin's (1994) study of ethnozoological nomenclature found evidence for size-symbolic naming schemes in birds across four languages, the same naming scheme was only observed in one language when applied to fish.

Symbolism in consonants

While there is a strong tradition focusing on the role of vowels in potential size-sound symbolism, the role of consonants has received comparatively little attention. Recently, however, a number of works have investigated the relationship. For instance, Klink's (2000) investigations of size-sound symbolism in company brand names suggested that names non-arbitrarily communicate meaningful information about the product (including size), with fricatives and voiceless obstruents being more readily associated with smaller products and stops and voiced obstruents more readily associated with larger ones. Indeed, in their study of size-symbolic name attributions to Pokémon characters, Kawahara and Kumagai (2019) observed an effect of voiced obstruents, such that voicing predicted size

in both native Japanese speakers and native English speakers – although less so in the latter case.

In an attempted replication the study, [Godoy et al. \(2020\)](#) explored the role of voiced obstruents in size-symbolic name attributions by native speakers of Brazilian Portuguese but did not observe the exact effect. In native speakers of Brazilian Portuguese, the effect of voiced obstruents (ready name attribution to post-evolution Pokémon characters) was more likely to occur when two voiced obstruents were present in a word (unlike the *single* voiced obstruent required for the same effect to be observed in native Japanese speakers; [Kawahara and Kumagai, 2019](#)) again suggesting learned language-specific differences in symbolism. Research on non-voiced consonants is much rarer, though results by [Winter and Perlman \(2021\)](#) suggested that [t], a voiceless stop, was indicative of smaller size. Finally, [Levickij \(2013\)](#) observed extensive correlations between phonemes symbolizing size in languages from 12 language families, such that close articulations were associated with smaller size and voiceless articulations were associated with smaller size and lighter weight.

Further, there are again several examples of interlinguistic inconsistencies across the size-symbolic literature, as applied to consonants specifically. For example, while [Duduciu and Ivan \(2015\)](#) replicated Klink's findings on iconicity in brand name vowels, they failed to observe the same effects of fricatives (as predicting of smaller size, vs. voiceless stops) in a Romanian sample. The same finding was also not corroborated by [Winter and Perlman \(2021\)](#). Moreover, voiced obstruents – low-frequency consonantal phonemes such as [dʒ] (i.e., the /j/ in the English “job”), assumed symbolic of greater size ([Ohala, 1984](#)) – has been shown similarly related to largeness in Chinese, English and Japanese – but not in Korean ([Shinohara and Kawahara, 2016](#)). Finally, [Saji et al. \(2019\)](#) found that native Japanese speakers tended to invent words involving voiced consonants for “big” and “heavy” moving objects, and words involving voiceless consonants for “small” and “light” ones – but the same relationship was not observed in native English speakers.

Human speech is, however, not made up solely of f_0 , vowels, and consonants, and a variety of voice cues including voice quality, and phonemic length and position, have recently been shown to facilitate size attributions.

Symbolism in other sources

Phonation type

[Akita \(2021\)](#) investigated iconicity in four phonation types – modal voice, creaky voice, falsetto, and whispering – in a sample of native Japanese speakers and found that voice creakiness was associated with larger images and whispering with smaller images. Creakiness reflects the voluntary recruitment and pulling together of laryngeal arytenoid cartilages, relaxing the vocal folds and slowing glottal air flow; creakiness is

also typically characterized by comparatively low f_0 ([Gordon and Ladefoged, 2001](#)). In comparison, voiceless sounds (i.e., whispering) are characteristically high-frequency signals. To the knowledge of the author, however, neither finding has yet been replicated in other languages.

Mora count

Recent work has also demonstrated a “longer-is-stronger” effect, such that larger Pokémon characters were attributed names with greater mora counts (i.e., long vowels; [Kawahara et al., 2018, 2020](#); [Kawahara, 2020](#)). Following up from these findings, [Kawahara and Kumagai \(2019\)](#) showed how both native Japanese and English speakers attributed names containing greater mora counts to evolved (i.e., larger) characters, compared to their un-evolved (smaller) forms (see also [Kawahara and Breiss, 2021](#)). It has been suggested this effect is comparable to quantitative iconicity in grammar ([Haiman, 1980](#); see discussion in [Kawahara et al., 2018](#); [Kawahara and Breiss, 2021](#)). If so, perception of extended vowel production (with greater mora count) more readily affords iconicity, compared to shorter productions.

Phoneme position

Finally, initial evidence of a phonemic positional effect was provided by [Kawahara et al. \(2008\)](#), who pointed to word-initial syllables as determinants of iconicity. Additionally, in a larger-scale study, [Haynie et al. \(2014\)](#) examined 120 Australian languages, observing significant positional effects. The authors found a higher occurrence of palatal consonants (e.g., [j], the /y/ in the English “yes”) in medial positions for “small” items, while low sounds were frequently associated with “large” items, regardless of position. Word-initial syllables are indeed meaningful for word recognition ([Marslen-Wilson, 1980](#)), but – unlike voice quality and length of phonation – positional effects are not obviously afforded by existing theories of size-sound symbolism. Explicating this issue should be the goal of future theoretical work on the topic.

Discussion

What explains size-sound symbolism?

While extensive research has focused on the phenomenon of size symbolism in speech, explanatory models have largely been neglected in comparison ([Lockwood and Dingemanse, 2015](#); [Sidhu and Pexman, 2018](#)). In this section, a set of relevant theories are briefly reviewed and discussed, with a focus on whether a given theory can explain (1) phonetic size-sound symbolic associations as a cognitive phenomenon (the Association criterion); and (2) inconsistent

effects of iconicity across studies of different languages (the Inconsistency criterion).

The frequency code

No discussion of phonetic size symbolism can be complete without explicit coverage of Ohala's (1984) frequency code hypothesis (FC). Indeed, evidence from bioacoustics has been provided based on research on vocalizations by range of mammalian species (for a review, see Charlton and Reby, 2016), including humans (Pisanski et al., 2014). Human cross-species size attribution has also been explored (Taylor et al., 2008). The relative contextual success of the theory stems from its reflecting a fundamental principle of physics, with the size of laryngeal component vibrating at slower rates (Titze, 1994). Nevertheless, theories purely based on bioacoustics have no ready explanations for the apparent interlinguistic differences commonly observed across studies: if associations were a universal, we should expect equivalent findings across languages. Thus, FC constitutes a strong candidate for meeting the Association criterion but does not meet the Inconsistency criterion.

Sound symbolism bootstrapping

The "Sound symbolism bootstrapping" hypothesis was presented by Imai and Kita (2014), who argued that symbolism helps infants establish associate speech sounds with referents. Cross-modal sound-symbolism correspondence does indeed appear early in life (Ozturk et al., 2013; Imai et al., 2015), facilitating word learning (Lockwood et al., 2016) and verb learning (Yoshida, 2012). Further, semantics of an image (as opposed to *de-facto* size) appear to drive symbolism (Auracher, 2017), suggesting the connection is, at least partially, learned. The bootstrapping hypothesis thus appears to complement FC, providing a potential developmental component that may facilitate subsequent mapping, partially meeting the Association criterion; however, it does contribute toward meeting the Inconsistency criterion.

Theories of embodiment

Embodied cognition (EC) accounts assume that aspects of cognition and information processing are contingent on bodily aspects, including the motor and perceptual systems (Lakoff and Johnson, 1980; Johnson, 1987). In speech research, such theories corroborate the motor theory of speech perception proposed by Liberman and colleagues (Liberman et al., 1967; Liberman and Mattingly, 1985), which argues that perception of speech is contingent on speech gestures. Intriguingly, a meta-analysis on the emergence of sound symbolism by Fort et al. (2018) suggest that a "bouba" effect (round sound-shape symbolism) emerges prior to a "kiki" effect (sharp sound-shape symbolism), in human infants². At birth, an

infant does have fully developed lips (i.e., for suckling) but cannot articulate plosives such as /k/, which require significant lingual development (Lieberman, 2012). Nevertheless, EC perspectives cannot explain why some languages appear to map iconic-phonetic features in contradictory ways (Diffloth, 1994; Inconsistency criterion).

Sapir-whorf hypotheses

Phonetic symbolism shares common origin with Sapir-Whorf hypotheses (SW) – assumptions that aspects of languages shape speakers' patterns of cognition (Sapir, 1929). While controversial in the literature (Deutscher, 2010; McWhorter, 2014), SW holds apparent value as potentially helping explore differences in symbolism across languages. For example, a probabilistic inference model of color perception by Cibelli et al. (2016) suggests that language-specific categories most significantly influence perception in conditions of uncertainty. There also appears to be direct exposure effects on pitch perception (Eitan, 2013), and possibly interacting effects of exposure on perception of auditory-spatial relationships (Ekström et al., 2022). However, while SW may provide an ostensible mechanistic explanation of interlinguistic inconsistencies (that is, why they exist at all), it does not present a reasonable explanation as to how inconsistencies arise (Association criterion); nor does it purport to explain the relationship between phonetic-symbolic features of spoken language (Inconsistency criterion).

Stochastic drift

Levickij (2013, p. 88) observed: "In a particular language [...] the potential synaesthetic effect will be weakened or come to nothing by phonetic and semantic laws of this language. The set of subjective symbolism should be therefore theoretically always larger than that of objective symbolism universals." Work comparing apparent phonetic symbolism across a great number of languages (Levickij, 2013; Haynie et al., 2014) may indeed be indicative of an explanatory role for randomness in interlinguistic differences in sound-symbolic sensitivity.

In population genetics, stochastic drift is the change of the average value of a random process (Fisher, 1922, 1930; Haldane, 1927). Drift is fundamental to evolutionary biology, explaining the appearance of novel traits and species as fixations of changes in sets of genotypes, resulting from changes in frequency across generations. Because linguistic forms are comparably copied between speakers and groups across generations, Newberry et al. (2017) argued, based on quantification of selection pressures

² Note, however, that judgements of size and shape appear to correspond to different acoustic cues, such that size is inferred from vowel F₁, F₂, and shape from F₂ and F₃ (Knoeferle et al., 2017; Saji et al., 2019; but see Elsen et al., 2021).

TABLE 1 Evaluation of five candidate theories and frameworks.

Theory	Meets Association criterion	Meets Inconsistency criterion
Frequency code hypothesis	✓	✗
Embodied cognition	✓	✗
Sound-symbolism Bootstrapping	✓	✗
Sapir-Whorf hypotheses	✗	✗
Stochastic drift	✗	✓

relative to stochastic drift in grammatical changes in English, that drift is likely stronger for irregular forms of (English) past-tense verbs and rare words. While such an approach may explain the evolution and spread of interlinguistic differences in size-symbolic cognitions (Inconsistency criterion; indeed, it is the only explanation discussed here that seemingly may do so), it does not lend itself to explaining the source of those cognitions themselves (Association criterion).

Concluding comments

In summary, while there have been extensive empirical research efforts targeting various aspects of size-sound symbolism, no cohesive theoretical framework yet exists that lends itself to explaining the nuanced findings observed. Here, two criteria were posited, but no theory discussed readily meets both (see Table 1).

The phenomenon of phonetic size-sound symbolism likely results from a combination of factors, and while findings may well be suggestive of some particular theory, theorists must seek independent motivation and validation. The challenge for the future phonetic symbologist, then, is independently motivating a theory that both accommodate evolutionary-biological principles of bioacoustics, while simultaneously accounting for interlinguistic differences. Reflecting the need for such a theory, future directions are briefly discussed in the final section.

Future directions

While a long-running tradition have utilized synthetic words, exaggerating effects of iconicity (Köhler, 1929;

Ramachandran and Hubbard, 2001; Westbury, 2005; Aveyard, 2012), more recent developments in the field have moved toward more realistic and integrative experiments. When stimulus words are derived from natural languages, iconicity effects are smaller compared to synthetic ones (Lockwood and Tuomainen, 2015; Styles and Gawne, 2017), but provide higher-validity data (e.g., Kanero et al., 2014). New research increasingly further points to a diversity of cues as responsible for sound-symbolic interpretations (Thompson and Estes, 2011; Knoeferle et al., 2017; Westbury et al., 2018), suggesting a more complex phenomenon than previously believed (e.g., Sapir, 1929). Importantly, there are observable individual differences in sound-symbolic sensitivity (Lockwood et al., 2016). To date, however, while a range of studies have observed group differences across native speakers of various languages, there has been no attempt known to the author, at isolating potential population-level differences in sound-symbolic sensitivity. Were such effects to be observed, they should provide valuable information toward proper understanding of inconsistencies observed in size-sound symbolism research, as well as contribute toward a cohesive theoretical framework that explicates them.

Author contributions

AE: conceptualization, investigation, writing—original draft, and writing—review and editing.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Akita, K. (2021). Phonation types matter in sound symbolism. *Cognit. Sci.* 45, e12982. doi: 10.1111/cogs.12982

Altmann, G. T. (1998). Ambiguity in sentence processing. *Trends Cognit. Sci.* 2, 146–152. doi: 10.1016/S1364-6613(98)01153-X

- Auracher, J. (2017). Sound iconicity of abstract concepts: Place of articulation is implicitly associated with abstract concepts of size and social dominance. *PLoS One* 12, e0187196. doi: 10.1371/journal.pone.0187196
- Aveyard, M. E. (2012). Some consonants sound curvy: Effects of sound symbolism on object recognition. *Memory Cognit.* 40, 83–92. doi: 10.3758/s13421-011-0139-3
- Berlin, B. (1994). Evidence for pervasive synesthetic sound symbolism in ethnozoological nomenclature. *Sound Symbolism* 76–93. doi: 10.1017/CBO9780511751806.006
- Bruckert, L., Liénard, J. S., Lacroix, A., Kreutzer, M., and Leboucher, G. (2006). Women use voice parameters to assess men's characteristics. *Proc. Royal Soc. Biol. Sci.* 273, 83–89. doi: 10.1098/rspb.2005.3265
- Charlton, B. D., and Reby, D. (2016). The evolution of acoustic size exaggeration in terrestrial mammals. *Nat. Commun.* 7, 1–8. doi: 10.1038/ncomms12739
- Cibelli, E., Xu, Y., Austerweil, J. L., Griffiths, T. L., and Regier, T. (2016). The Sapir-Whorf hypothesis and probabilistic inference: Evidence from the domain of color. *PLoS One* 11, e0158725. doi: 10.1371/journal.pone.0158725
- Delgado, R. A. (2006). Sexual selection in the loud calls of male primates: signal content and function. *Int. J. Primatol.* 27, 5–25. doi: 10.1007/s10764-005-9001-4
- Deutscher, G. (2010). *Through the Language Glass: Why the World Looks Different in Other Languages*. New York, NY: Metropolitan Books.
- Diffloth, G. (1994). "I: big, a: small," in *Sound Symbolism*, eds L. Hinton, J. Nichols, and J. Ohala (Cambridge: Cambridge University Press), 107–114.
- Duduciuc, A. C., and Ivan, L. (2015). *Brand Naming: Sound Symbolism, Brand Preference and Brand Performance. Studies and Scientific Researches. Economics Edition*. New Haven: Yale University Press
- Eitan, Z. (2013). How pitch and loudness shape musical space and motion. *Psychol. Music Multimedia* 27, 165–191. doi: 10.1093/acprof:oso/9780199608157.003.0008
- Ekström, A. G., Nirme, J., and Gärdenfors, P. (2022). Motion iconicity in prosody. *Front. Commun.* 7, 994162. doi: 10.3389/fcomm.2022.994162
- Elsen, H., Németh, R., and Kovács, L. (2021). The sound of size revisited—New insights from a German-Hungarian comparative study on sound symbolism. *Lang. Sci.* 85, 101360. doi: 10.1016/j.langsci.2021.101360
- Ey, E., Pfefferle, D., and Fischer, J. (2007). Do age- and sex-related variations reliably reflect body size in non-human primate vocalizations? A review. *Primates* 48, 253–267. doi: 10.1007/s10329-006-0033-y
- Fant, G. (1960). *The Acoustic Theory of Speech Production*. Mouton: The Hague.
- Fisher, R. A. (1922). On the dominance ratio. *Proc. Royal Soc. Edinburgh* 42, 321–341. doi: 10.1017/S0370164600023993
- Fisher, R. A. (1930). *The Genetical Theory of Natural Selection*. Clarendon Press: Oxford.
- Fitch, W. T. (1997). Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *J. Acoustical Soc. Am.* 102, 1213–1222. doi: 10.1121/1.421048
- Fort, M., Lammertink, I., and Peperkamp, S., Guevara-Rukoz, A., Fikkert, P., and Tsuji, S. (2018). Symbouki: a meta-analysis on the emergence of sound symbolism in early language acquisition. *Dev. Sci.* 21, e12659. doi: 10.1111/desc.12659
- Garcia, M., Herbst, C. T., Bowling, D. L., Dunn, J. C., and Fitch, W. (2017). Acoustic allometry revisited: morphological determinants of fundamental frequency in primate vocal production. *Sci. Rep.* 7, 1–11. doi: 10.1038/s41598-017-11000-x
- Godoy, M. C., Souza, De F. Filho Souza, N. S., França, J. G. M. H. A., and Kawahara, S. (2020). Gotta name'em all: An experimental study on the sound symbolism of Pokémon names in Brazilian Portuguese. *J. Psycholing. Res.* 49, 717–740. doi: 10.1007/s10936-019-09679-2
- González, J. (2004). Formant frequencies and body size of speaker: a weak relationship in adult humans. *J. Phonetics* 32, 277–287. doi: 10.1016/S0095-4470(03)00049-4
- Gordon, M., and Ladefoged, P. (2001). Phonation types: a cross-linguistic overview. *J. Phonetics* 29, 383–406. doi: 10.1006/jpho.2001.0147
- Grawunder, S., Crockford, C., Clay, Z., Kalan, A. K., Stevens, J. M., Stoessel, A., et al. (2018). Higher fundamental frequency in bonobos is explained by larynx morphology. *Curr. Biol.* 28, R1188–R1189. doi: 10.1016/j.cub.2018.09.030
- Gunter, C. D., and Manning, W. H. (1982). Listener estimations of speaker height and weight in unfiltered and filtered conditions. *J. Phonetics* 10, 251–257. doi: 10.1016/S0095-4470(19)30986-6
- Haiman, J. (1980). The iconicity of grammar: Isomorphism and motivation. *Language* 1, 515–540. doi: 10.2307/414448
- Haldane, J. B. S. (1927). A mathematical theory of natural and artificial selection, part V: selection and mutation. *Mathematical Pro. Cambridge Philos. Soc.* 23, 838–844. doi: 10.1017/S0305004100015644
- Hauser, M. D. (1993). The evolution of nonhuman primate vocalizations: effects of phylogeny, body weight, and social context. *Am. Nat.* 142, 528–542. doi: 10.1086/285553
- Haynie, H., and Bowern, C., and LaPalombara, H. (2014). Sound symbolism in the languages of Australia. *PLoS ONE* 9, e92852. doi: 10.1371/journal.pone.0092852
- Hoshi, H., Kwon, N., Akita, K., and Auracher, J. (2019). Semantic associations dominate over perceptual associations in vowel-size iconicity. *i-Perception* 10, 2041669519861981. doi: 10.1177/2041669519861981
- Imai, M., and Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Philos. Trans. Royal Soc. Biol. Sci.* 369, 20130298. doi: 10.1098/rstb.2013.0298
- Imai, M., Miyazaki, M., Yeung, H. H., Hidaka, S., Kantartzis, K., Okada, H., et al. (2015). Sound symbolism facilitates word learning in 14-month-olds. *PLoS One* 10, e0116494. doi: 10.1371/journal.pone.0116494
- Jespersen, O. (1933). Symbolic value of the vowel *i*. *Philologica* 1, 1–19.
- Johnson, M. (1987). *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. Chicago: University of Chicago Press.
- Kanero, J., Imai, M., Okuda, J., Okada, H., and Matsuda, T. (2014). How sound symbolism is processed in the brain: a study on Japanese mimetic words. *PLoS one* 9, e97905. doi: 10.1371/journal.pone.0097905
- Kawahara, S. (2020). A wug-shaped curve in sound symbolism: The case of Japanese Pokémon names. *Phonology* 37, 383–418. doi: 10.1017/S0952675720000202
- Kawahara, S., and Breiss, C. (2021). Exploring the nature of cumulativity in sound symbolism: experimental studies of Pokémon names with English speakers. *Lab. Phonol.* 12, 280. doi: 10.5334/labphon.280
- Kawahara, S., and Kumagai, G. (2019). Expressing evolution in Pokémon names: experimental explorations. *J. Japanese Ling.* 35, 3–38. doi: 10.1515/jjl-2019-2002
- Kawahara, S., Noto, A., and Kumagai, G. (2018). Sound symbolic patterns in Pokémon names. *Phonetica* 75, 219–244. doi: 10.1159/000484938
- Kawahara, S., Shinohara, K., and Uchimoto, Y. (2008). A positional effect in sound symbolism: An experimental study. *Proc. Japan Cognit. Ling. Assoc.* 8, 417–427.
- Kawahara, S., Suzuki, M., and Kumagai, G. (2020). *Sound symbolic patterns in Pokémon move names in Japanese. ICU Working Papers in Linguistics 10. Festschrift for Prof. Junko Hibiya in the occasion of her retirement from ICU*, 17–30.
- Klink, R. R. (2000). Creating brand names with meaning: the use of sound symbolism. *Marketing Letters* 11, 5–20. doi: 10.1023/A:1008184423824
- Knoefler, K., Li, J., Maggioni, E., and Spence, C. (2017). What drives sound symbolism? Different acoustic cues underlie sound-size and sound-shape mappings. *Sci. Rep.* 7, 1–11. doi: 10.1038/s41598-017-05965-y
- Köhler, W. (1929). *Gestalt Psychology*. New York, NY: Liveright.
- Küssner, M. B., Tidhar, D., Prior, H. M., and Leech-Wilkinson, D. (2014). Musicians are more consistent: Gestural cross-modal mappings of pitch, loudness and tempo in real-time. *Front. Psychol.* 5, 789. doi: 10.3389/fpsyg.2014.00789
- Lakoff, G., and Johnson, M. (1980). *Metaphors We Live By*. Chicago: Chicago University Press.
- Levickij, V. V. (2013). Phonetic symbolism in natural languages. *Glottology* 4, 72–91. doi: 10.1524/glott.2013.0006
- Lieberman, A. M., Cooper, F. S., Shankweiler, D. P., and Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychol. Rev.* 74, 431. doi: 10.1037/h0020279
- Lieberman, A. M., and Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition* 21, 1–36. doi: 10.1016/0010-0277(85)90021-6
- Lieberman, P. (2012). Vocal tract anatomy and the neural bases of talking. *J. Phonetics* 40, 608–622. doi: 10.1016/j.wocn.2012.04.001
- Lockwood, G., and Dingemans, M. (2015). Iconicity in the lab: A review of behavioral, developmental, and neuroimaging research into sound-symbolism. *Front. Psychol.* 6, 1246. doi: 10.3389/fpsyg.2015.01246
- Lockwood, G., Hagoort, P., and Dingemans, M. (2016). How iconicity helps people learn new words: neural correlates and individual differences in sound-symbolic bootstrapping. *Collabra* 2, 42. doi: 10.1525/collabra.42
- Lockwood, G., and Tuomainen, J. (2015). Ideophones in Japanese modulate the P2 and late positive complex responses. *Front. Psychol.* 6, 933. doi: 10.3389/fpsyg.2015.00933

- Marslen-Wilson, W. D. (1980). "Speech understanding as a psychological process" in *Spoken Language Generation and Understanding*. NATO Advanced Study Institutes Series, ed J. C. Simon (Dordrecht: Springer), 59. doi: 10.1007/978-94-009-9091-3_2
- Masataka, N. (1994). Lack of correlation between body size and frequency of vocalizations in young female Japanese macaques (*Macaca fuscata*). *Folia Primatol.* 63, 115–118. doi: 10.1159/000156803
- McComb, K. E. (1991). Female choice for high roaring rates in red deer, *Cervus elaphus*. *Anim. Behav.* 41, 79–88. doi: 10.1016/S0003-3472(05)80504-4
- McWhorter, J. H. (2014). *The Language Hoax: Why the World Looks the Same in Any Language*. Oxford: Oxford University Press.
- Newberry, M. G., Ahern, C. A., Clark, R., and Plotkin, J. B. (2017). Detecting evolutionary forces in language change. *Nature* 551, 223–226. doi: 10.1038/nature24455
- Newman, S. R., Butler, J., Hammond, E. H., and Gray, S. D. (2000). Preliminary report on hormone receptors in the human vocal fold. *J. Voice.* 14, 72–81. doi: 10.1016/S0892-1997(00)80096-x
- Ohala, J. J. (1984). An ethological perspective on common cross-language utilization of F0 of voice. *Phonetica* 41, 1–16. doi: 10.1159/000261706
- Ohala, J. J. (1994). The frequency code underlies the sound-symbolic use of voice pitch. *Sound Symbol.* 2, 325–347. doi: 10.1017/CBO9780511751806.022
- Owren, M. J., Seyfarth, R. M., and Cheney, D. L. (1997). The acoustic features of vowel-like grunt calls in chacma baboons (*Papio cyncephalus ursinus*): Implications for production processes and functions. *J. Acoust. Soc. Am.* 101, 2951–2963. doi: 10.1121/1.418523
- Ozturk, O., Krehm, M., and Vouloumanos, A. (2013). Sound symbolism in infancy: Evidence for sound–shape cross-modal correspondences in 4-month-olds. *J. Exp. Child Psychol.* 114, 173–186. doi: 10.1016/j.jecp.2012.05.004
- Peterfalvi, J. M. (1965). Les recherches expérimentales sur le symbolisme phonétique. *Lannée Psychol.* 65, 439–474. doi: 10.3406/psy.1965.27443
- Pisanski, K., Fraccaro, P. J., Tighe, C. C., O'Connor, J. J., Röder, S., Andrews, P. W., et al. (2014). Vocal indicators of body size in men and women: a meta-analysis. *Anim. Behav.* 95, 89–99. doi: 10.1016/j.anbehav.2014.06.011
- Puts, D. A., Hill, A. K., Bailey, D. H., Walker, R. S., Rendall, D., Wheatley, J. R., et al. (2016). Sexual selection on male vocal fundamental frequency in humans and other anthropoids. *Proc. Royal Soc. Biol. Sci.* 283, 20152830. doi: 10.1098/rspb.2015.2830
- Ramachandran, V. S., and Hubbard, E. M. (2001). Synaesthesia—a window into perception, thought and language. *J. Consciousness Stu.* 8, 3–34.
- Riede, T., and Brown, C. (2013). Body size, vocal fold length, and fundamental frequency—implications for mammal vocal communication. *Nova Acta Leopoldina NF* 111, 1–20.
- Saji, N., Akita, K., Kantartzis, K., Kita, S., and Imai, M. (2019). Cross-linguistically shared and language-specific sound symbolism in novel words elicited by locomotion videos in Japanese and English. *PLoS One* 14, e0218707. doi: 10.1371/journal.pone.0218707
- Sapir, E. (1929). A study in phonetic symbolism. *J. Exp. Psychol.* 12, 225. doi: 10.1037/h0070931
- Shinohara, K., and Kawahara, S. (2010). A cross-linguistic study of sound symbolism: The images of size. *Ann. Meeting Berkeley Ling. Soc.* 36, 396–410. doi: 10.3765/bls.v36i1.3926
- Shinohara, K., and Kawahara, S. (2016). *A Cross-Linguistic Study of Sound Symbolism: The Images of Size. Paper Presented at: Proceedings of the Thirty Sixth Annual Meeting of the Berkeley Linguistics Society*. Berkeley Linguistics Society, Berkeley, 396–410.
- Shinohara, K., Kawahara, S., and Tanaka, H. (2020). Visual and proprioceptive perceptions evoke motion-sound symbolism: Different acceleration profiles are associated with different types of consonants. *Front. Psychol.* 11, 589797. doi: 10.3389/fpsyg.2020.589797
- Shintel, H., Nusbaum, H. C., and Okrent, A. (2006). Analog acoustic expression in speech communication. *J. Memory Lang.* 55, 167–177. doi: 10.1016/j.jml.2006.03.002
- Sidhu, D. M., and Pexman, P. M. (2018). Five mechanisms of sound symbolic association. *Psycho. Bullet. Rev.* 25, 1619–1643. doi: 10.3758/s13423-017-1361-1
- Styles, S. J., and Gawne, L. (2017). When does maluma/takete fail? Two key failures and a meta-analysis suggest that phonology and phonotactics matter. *i-Perception* 8, 2041669517724807. doi: 10.1177/2041669517724807
- Tarte, R. D. (1974). Phonetic symbolism in adult native speakers of Czech. *Lang. Speech* 17, 87–94. doi: 10.1177/002383097401700109
- Taylor, A. M., Reby, D., and McComb, K. (2008). Human listeners attend to size information in domestic dog growls. *J. Acoustical Soc. Am.* 123, 2903–2909. doi: 10.1121/1.2896962
- Thompson, P. D., and Estes, Z. (2011). Sound symbolic naming of novel objects is a graded function. *Q. J. Exp. Psychol.* 64, 2392–2404. doi: 10.1080/17470218.2011.605898
- Titze, I. R. (1994). *Principles of Voice Production*. Englewood Cliffs, NJ: Prentice Hall.
- Tsur, R. (2006). Size–sound symbolism revisited. *J. Pragmatics* 38, 905–924. doi: 10.1016/j.pragma.2005.12.002
- Utan, R. (1978). Size-sound symbolism. *Univ. Hum. Lang.* 2, 525–568.
- Wermke, K., and Robb, M. P. (2010). Fundamental frequency of neonatal crying: Does body size matter?. *J. Voice* 24, 388–394. doi: 10.1016/j.jvoice.2008.11.002
- Westbury, C. (2005). Implicit sound symbolism in lexical access: evidence from an interference task. *Brain Lang.* 93, 10–19. doi: 10.1016/j.bandl.2004.07.006
- Westbury, C., Hollis, G., Sidhu, D. M., and Pexman, P. M. (2018). Weighing up the evidence for sound symbolism: distributional properties predict cue strength. *J. Memory Lang.* 99, 122–150. doi: 10.1016/j.jml.2017.09.006
- Winter, B., and Perlman, M. (2021). Size sound symbolism in the English lexicon. *Glossa J. General Ling.* 6:1646. doi: 10.5334/gjgl.1646
- Yoshida, H. (2012). A cross-linguistic study of sound symbolism in children's verb learning. *J. Cognit. Dev.* 13, 232–265. doi: 10.1080/15248372.2011.573515