Language and Linguistics Compass

WILEY

REVIEW ARTICLE OPEN ACCESS

Sound Symbolism in the Lexicon: A Review of Iconic-Systematicity

David M. Sidhu

Department of Psychology, Carleton University, Ottawa, Canada

Correspondence: David M. Sidhu (david.sidhu@carleton.ca)

Received: 27 August 2024 | **Revised:** 8 November 2024 | **Accepted:** 11 November 2024

Keywords: bouba/kiki effect | iconic-systematicity | iconicity | language evolution | maluma takete effect | mil/mal effect | non-arbitrariness | sound symbolism | systematicity

ABSTRACT

Sound symbolism refers to associations between language sounds (i.e., phonemes) and particular properties (e.g., certain shapes). For example, phonemes like /m/ are associated with roundness, while phonemes like /k/ are associated with spikiness. In this paper I review the accumulating evidence that different instances of sound symbolism can be observed as patterns in real words in existing lexicons (e.g., /m/ occurring more frequently in words for round things). The properties examined include shape, size, texture, valence and arousal. Such effects are an instance of both iconicity (words whose forms resemble their meanings) and systematicity (largescale patterns in the forms of related words in a language). I also discuss open questions on the topic, including how such patterns emerge, and their effects on language processing.

1 | Introduction

Sound symbolism refers to associations evoked by the sounds of language (i.e., *phonemes*). For instance, participants report that "maluma" is a more appropriate label for a round shape, while "takete" is a more appropriate label for a spiky shape (Ćwiek et al. 2022; Köhler 1929; A. Nielsen and Rendall 2011; Sidhu and Pexman 2017; Styles and Gawne 2017). This reveals an association between the phonemes in either pseudoword and certain kinds of shapes. The association exists for individuals from different languages, cultures, and writing systems (Ćwiek et al. 2022; Styles and Gawne 2017; also see these papers for exceptions); as well as across ages (Fort et al. 2018). In another example, participants report that /i/ is more appropriate for small sizes while /ɑ/ is more appropriate for large sizes (Knoeferle et al. 2017; Sapir 1929; Thompson and Estes 2011). Associations have also been demonstrated for dimensions such as brightness (Newman 1933), speed (Cuskley 2013), emotion (Aryani et al. 2018; Yu, McBeath, and Glenberg 2021), personality (Sidhu et al. 2019), and others (Sidhu, Vigliocco, and Pexman 2022; for reviews see Lockwood and Dingemanse 2015; Sidhu and Pexman 2018). These effects are often explained in terms of crossmodal analogies between one of the multimodal features of phonemes (e.g., sound, kinesthetics, mouth shape) and the associated dimension (for a review, see Sidhu and Pexman 2018).

The previously reviewed work has examined sound symbolic associations using pseudowords. In this paper I review evidence that these associations exist as patterns in real words, in existing lexicons. For example, that words for small things are more likely than chance to contain phonemes that individuals associate with smallness. This question has been present in work on sound symbolism from the very beginning. In his pioneering paper on size sound symbolism, Sapir (1929) suggested that:

> It would be an important check to amass a large number of randomly distributed meaningful words, to classify into the two groups of 'large' and 'small'...and

> to see if...the distributions are of the same nature as those studied in the experiments.
>
> (p. 235)

1.1 | Iconicity, Systematicity and Iconic-Systematicity

A word for "small" containing phonemes associated with smallness would represent an instance of *iconicity*. This refers to signals for which there is a sense of resemblance between some aspect of form (e.g., sound, feeling of articulation) and some aspect of meaning (see Dingemanse et al. 2015; Occhino et al. 2017; Perniss, Thompson, and Vigliocco 2010; Winter, Woodin, and Perlman 2023).[1] While some instances of iconicity in spoken language involve imitation of a sound (e.g., onomatopoeia like "whoosh" or "oink"), our interest is in what some have referred to as *crossmodal iconicity*: when the sound of a word resembles a property from a different modality (e.g., sound resembling size).

Importantly, it would not be enough to point out that a word like "teeny" contains the small-associated phoneme /i/. In order to make the claim that crossmodal iconicity exists as a pattern in the lexicon, researchers would need to show that this pattern appears more often than would be expected by chance. That is, researchers need a distribution of words on which to conduct statistical tests. There have been two approaches to amassing such a distribution: to look for a pattern in translations of the word "small" *across* languages, or to look for a pattern in words for smallness, or small things, *within* a language. Here I review the latter approach.[2]

The term for statistical patterns in the forms of related words in a language is *systematicity* (Dingemanse et al. 2015). For example, researchers have demonstrated systematic differences in the forms of nouns and verbs (Farmer, Christiansen, and Monaghan 2006), and concrete and abstract words (Reilly et al. 2012). Crucially, iconicity and systematicity are orthogonal properties. A systematic pattern need not be iconic. Consider a hypothetical example in which words for small animals tended to contain the phoneme /h/. It would be difficult to make the case that this pattern was iconic—that the phoneme /h/ somehow resembled smallness. This would be an instance of systematicity alone. However, if it were the case that words for small animals tended to contain the phoneme /i/—a sound that individuals associate with smallness—this would be an instance of *iconic-systematicity*. This is distinct from iconicity alone as it appears as a statistical pattern in the lexicon. Iconic-systematicity is the focus of the present review.

Iconic-systematicity is probabilistic rather than deterministic. That is, we would not expect all words for small animals to contain the phoneme /i/. Instead, we would expect the phoneme /i/ to appear in these words more often than expected by chance alone. There are many factors that affect language change, and a bias towards iconic forms would only be one of them. Indeed, there are also factors which work *against* iconicity. A fully iconic vocabulary would reduce discriminability among related words (see Gasser 2004; Monaghan, Mattock, and Walker 2012). Such a language could also have difficulty expressing certain concepts which are difficult for sounds to imitate (Winter et al. 2017).

In the sections that follow I outline the general approach that researchers have taken in searching for iconic-systematicity, review some of the most notable demonstrations and discuss several open questions.

2 | Methods for Demonstrating Iconic-Systematicity

There have been two general approaches to demonstrating iconic-systematicity: a *categorical approach* and a *continuous approach*. The categorical approach is to generate lists of words that are relevant to some meaning contrast. For example, a list of adjectives describing smallness and a list of adjectives describing largeness. A common method has been to use collect synonyms for relevant words (e.g., synonyms for "tiny" and "small"; Winter and Perlman 2021). Researchers then examine whether any phonemes are more likely to be present in one category than the other (e.g., Winter and Perlman 2021). In a continuous approach, researchers analyse a set of words that have been rated on some meaning dimension (e.g., their emotional arousal; Aryani et al. 2018). Sometimes this is restricted to certain kinds of words (e.g., object nouns; Sidhu et al. 2021), and sometimes not (e.g., Aryani et al. 2018). Researchers then examine whether the presence of a given phoneme predicts higher or lower ratings on that dimension. An advantage of the categorical approach is that word meanings are very specific to the relevant contrast. An advantage of the continuous approach is that it allows the inclusion of a greater number of items, with more diverse meanings (e.g., objects of all sizes as opposed to those at the extremes; easily accommodating many parts of speech).

With either approach, researchers must test whether phonemes are predictive of meaning (i.e., category membership or rating). This can happen in a bottom-up exploratory, or top-down confirmatory approach. In a bottom-up approach, all phonemes are included as candidates and researchers investigate if any are significant predictors. These tests must be able to deal with many correlated predictors, leading to the use of advanced statistical approaches such as a random forest approach (see Strobl et al. 2007). In a top-down approach researchers begin with a priori hypotheses about which phonemes are relevant to a particular meaning contrast, and then test effects of those specific phonemes.

No matter the approach, researchers must have a way to establish that a particular phoneme-meaning relationship is iconic. This can be challenging as iconicity is by its very nature a subjective phenomenon that exists in the eye of the beholder (Occhino et al. 2017; for a discussion of how to quantify iconicity see Dingemanse, Perlman, and Perniss 2020). That is, one person may have the sense that /i/ resembles smallness, while another may not. The most common approach has been to use the results of studies in which pseudowords were rated on the relevant dimension. Because raters have little semantic

information available for ratings, this ostensibly serves to isolate the crossmodal associations of a given phoneme. Researchers might also make a theoretical argument that a certain phoneme resembles the meaning in question (e.g., suggesting that articulating a certain phoneme overlaps with the expression of a certain emotion; Yu, McBeath, and Glenberg 2021).

It is notable that most of the studies to follow include an analysis on monomorphemic items. This is because iconic-systematicity could arise from the pervasiveness of an iconic affix. For example, the suffix "-y" (/i/) can be used to indicate a small size (e.g., "doggy" from "dog"; Trask 2000). This could lead to a pattern in which words containing the phoneme /i/ are smaller in size. This is a different phenomenon from one in which a variety of roots related to smallness contain the phoneme /i/.

3 | Demonstrations of Iconic-Systematicity

3.1 | Shape

Shape sound symbolism refers to an association of back rounded vowels (e.g., /o/), sonorants (e.g., /l, m/), and less consistently voiced stops (e.g., /b, g/) with round curvy shapes; and of front unrounded vowels (e.g., /i/) and voiceless stops (e.g., /t, k/) with jagged spiky shapes (i.e., the maluma/takete or bouba/kiki effect; D'Onofrio 2014; Köhler 1929; Ramachandran and Hubbard 2001; Sidhu, Vigliocco, and Pexman 2022; Westbury et al. 2018). In an early exploration of this effect in language, Katz (1986) analysed 64 concrete words (i.e., words for things that can be experienced with the senses) which had been rated on their shape. Note that unless otherwise stated, the studies to be reviewed include English words. These words were monosyllabic and only contained one vowel. Katz analysed words based on the vowel letter they contained. The only effect to emerge was that words containing the letter "U" were rounder than those containing an "E". A drawback of this study is that it was based on spelling rather than sound. More recently, Monaghan, Mattock, and Walker (2012) analysed 509 adjectives conveying roundness or angularity. They found that velar (e.g., /g, k/) and voiceless consonants (e.g., /t, k/) were more likely to occur in words for angularity. However, the authors noted that these effects were no longer significant when correcting for multiple comparisons (i.e., the separate analyses for each phoneme feature).

Sidhu et al. (2021) took a large-scale approach to this question by examining the phonemes appearing in 1,757 English object words. They first collected data quantifying the shapes of the objects that each word referred to. Then, they examined whether the presence of any English phonemes were predictive of a word referent's shape. They found that /u, m, oʊ, b, i/ were predictive of a rounder shape, while /aɪ, tʃ, k, ʃ, ɝ, r, t, ɪ, s/ were predictive of a spikier shape (see Table 1 for a summary of select key findings in the papers discussed). Notably, nearly all of these effects were consistent with associations of these phonemes in pseudoword studies, supporting this as an instance of iconic-systematicity. This was confirmed in a subsequent analysis which first coded each word based on the shape associations of its component phonemes (based on previous findings). The authors found that each word's sound association was indeed predictive of the shape of the object to which it referred.

3.2 | Size

The other longstanding sound symbolism effect is size. Individuals associate front and high vowels (e.g.,/i/) with smallness, and low and back vowels (e.g., /ɑ/) with largeness (i.e., the mil/mal effect; Knoeferle et al. 2017; Sapir 1929; Thompson and Estes 2011). Voiceless stops (e.g.,/t, k/) also appear to be associated with smallness, while voiced stops (e.g., /d, g/) and sonorants (e.g., /l, m/) are associated with largeness (Knoeferle et al. 2017).

Writing even before Sapir's pioneering paper on size sound symbolism[3], Jespersen (1922) accumulated illustrative lists of English words for "little", children, young animals, small things, and shortening, which contained the phoneme /i/. In an early statistical investigation of this pattern, Newman (1933) used a thesaurus to collect 351 words denoting smallness or largeness. He determined that the phonemes appearing in either group did not differ in their size associations. However, it should be noted that some word selections were questionable (e.g., "spare" as a word denoting smallness).

In a later study, the classification of words' denotations was shifted to participants. Johnson (1967) gave participants 3 minutes to produce as many words for large and small as they could (324 in total). Words were classified based on their first vowel sound, and Johnson calculated how often each vowel phoneme appeared first in a large versus small word. There was a correlation between each phoneme's appearance in large versus small words, and its size association. Interestingly, a similar analysis of consonants found no effect. The previously mentioned study by Katz (1986) also examined a set of 60 concrete words that had been rated in terms of their size. He did not find any evidence of iconic-systematicity. Berlin (1992) examined the sounds in names for animals of different sizes. In one such study on Huambisa, he found that /i/ was especially common in the names of birds smaller than 10 inches, while /e, a, u/ were more common in names of larger birds.

Recently, Winter and Perlman (2021) explored this question in two ways. They first analysed 52 English synonyms of "small" and "large" to determine if the presence of any phonemes were predictive of a word's meaning. Indeed, they found that /i, ɪ, t/ were predictive of a small adjective, while /ɑ/ was predictive of a large adjective. Each of these effects are consistent with phonemes' size associations in pseudoword studies. Interestingly, they next analysed 4683 words from throughout the lexicon (i.e., beyond size adjectives). No phonemes were predictive of size in this larger more general sample. This is consistent with Sidhu et al. (2021), who found that no phonemes predicted the sizes of their object words ($n = 613$ words with available size ratings). A key difference here may be that size is central to the meanings of adjectives such as "little" and "massive". Conversely, size is only one feature of words such as "pin" or "elephant". These results help clarify that some instances of iconic-systematicity are restricted to particular parts of the lexicon.

TABLE 1 | Select key findings from recent papers.

Study	Dimension	Key finding	Words analysed
Sidhu et al. (2021)	Shape	/u, m, oʊ, b, i/ more common in rounder objects; /aɪ, tʃ, *k*, ʃ, ɝ, r, t, ɪ, s/ more common in spikier objects	1757 English object nouns
Winter and Perlman (2021)	Size	/i, ɪ, t/ more common in words for "small"; /ɑ/ more common in words for "large"	52 English words for "small" or "large"
Winter et al. (2022)	Texture	/r/ more common in words for "rough"[a]	681 words from 332 languages for "rough" or "smooth". /r/ was the only phoneme tested in this particular analysis.
Yu, McBeath, and Glenberg (2021)	Valence	/i/ more common in pleasant words; /ʌ/ more common in unpleasant words	3329 English words containing only an /i/ or /ʌ/. This was the only comparison tested.
Aryani et al. (2018)	Arousal	Short vowels, voiceless consonants and stop consonants more common in exciting words; long vowels and voiced consonants more common in calming words.	2694 German words

Note: Each study cited contains more findings than those included in the table.
[a]Particularly in languages with a trilled /r/.

3.3 | Texture

This refers to the smoothness versus roughness of a given surface. Winter et al. (2022; see also Ćwiek et al. 2024) noted that the articulation of the phoneme /r/ resembles the tactile sensation of a rough surface texture. They investigated 99 English and 85 Hungarian adjectives that described a rough versus smooth surface texture. They found that /r/ was indeed more common in rough versus smooth words. This pattern was also shown cross-linguistically in words for "rough" and "smooth" across 332 languages. Notably, this was particular to languages with a trilled realization of /r/ (i.e., a "rolled r"), an articulation that heightens the resemblance. Finally, the authors even examined Proto-Indo-European: the reconstructed language which served as the ancestor to all Indo-European languages, spoken between 5500 and 8000 years ago (Mallory and Adams 2006). Winter et al. examined 39 reconstructed root words related to roughness and smoothness and found that /r/ was significantly more common in the roots related to roughness. This indicated that the observed iconic-systematicity in present day languages may have its origins thousands of years in the past.

3.4 | Emotion

Emotion is typically categorised along two axes: valence (how unpleasant vs. pleasant something is) and arousal (how calming vs. exciting something is). There have been studies examining iconic-systematicity for both.

3.4.1 | Valence

In an early attempt at quantifying valence iconicity in language, Thorndike (1945) developed a list of pleasant and unpleasant words in English. These included words for beauty versus ugliness, sweet versus sour flavours, soft versus prickly textures, and so on. He compared the occurrence of each phoneme in pleasant versus unpleasant words and found that some were more common in one group than the other (e.g., /ʒ/ was more common in pleasant words; /ʌ/ in unpleasant words). A downside of this study is that the stimuli were selected based on his intuition. However, it is notable that at least two other early papers also noted a link between the phoneme /ʌ/ and unpleasantness (Householder 1946; Jespersen 1922; as cited in Taylor and Taylor 1965). The previously mentioned study by Katz (1986) examined 86 monosyllabic words rated for their emotional content and did not find that the presence of any letter depended on valence.

More recently, Yu, McBeath, and Glenberg (2021) conducted a thorough examination of one specific phoneme contrast. They proposed that the phoneme /i/ should be associated with positive valence because its articulation involves contraction in the zygomaticus major muscle which is involved in smiling. Conversely articulating the phoneme /ʌ/ should be associated with negative valence because it is consistent with negative facial emotion. They examined the valence of 87 word pairs that differed only in the presence of /i/ and /ʌ/ (e.g., "gleam" and "glum"; i.e., *minimal pairs*). In 66% of the pairs, the word containing /i/ was the more positive of the two. They then also examined 3329 words with available valence ratings and containing either the phoneme /i/ or /ʌ/, and found that those containing /i/ were more positive.

Notably, this "gleam-glum effect" is observable in a previous study by Adelman, Estes, and Cossu (2018), who used an entirely bottom-up approach to examine which phonemes were predictive of words' valence, in five different languages. While many phonemes were significant predictors, it is notable that /i/ and /ʌ/ were positive and negative predictors, respectively. However, /i/ was no longer significant after removing multimorphemic words.

3.4.2 | Arousal

Katz (1986) discovered that, among the 86 monosyllabic words which had been rated on their emotion, those containing the

letter "I" were higher in arousal than those containing an "E". More recently, Aryani et al. (2018) took a bottom-up approach to exploring emotion iconicity in German. However, rather than focus on the presence of phonemes, their main interest was in words' underlying acoustic properties. To that end, they examined whether 11 different acoustic properties predicted both arousal and valence ratings of 2574 German words.[4] They discovered that more arousing words tended to have a higher first formant frequency (*F1*; indicative of vowels articulated lower in the mouth), greater variation in F1 and F3 (indicative of lip rounding) frequency, a lower overall intensity, and greater variation in overall frequencies.

The authors then further investigated whether these acoustic patterns were related to the presence of particular phonemes. This led to the discovery that more arousing words tended to have short vowels (e.g., /ɪ/ instead of /i/). Their interpretation was this this could be related to the short breaths taken when a person is in a high state of arousal. In addition, they found that the presence of voiceless obstruents, voiced/voiceless stop consonants, and hissing sibilants (e.g., /s/) were predictive of greater arousal. This latter sound is involved in making a high arousing sound to get a person's attention (e.g., "psst"). Thus, arousal represents a unique case in which phoneme associations may arise from the usage of sounds in particular emotional states.

3.5 | Phonesthemes

It is worth briefly discussing the topic of *phonesthemes*: sounds which frequently occur in related words (Bergen 2004). For example, "gl-" occurs as an onset in many words related to light (e.g., "glimmer", "glisten", "glow"). Phonesthemes involve a stricter requirement of phoneme location than the examples discussed so far. That is, a given phonestheme is typically defined as appearing at the onset of a syllable (e.g., "gl-") or as the rime (e.g., "-ash" in words for a collision; e.g., "crash", "smash"). They are also often made up of more than a single phoneme.

Phonesthemes are an excellent illustration of the difference between systematicity and iconic-systematicity. All phonesthemes are examples of systematicity in the lexicon. However, not all of these are iconic. It would be difficult to make the case for an iconic relationship between "gl-" and light. Conversely, the phonestheme "sn-", occurring at the beginning of words related to the nose (e.g., "sneeze", "snore", "sniff"), contains the nasal phoneme /n/. This supports "sn-" as an instance of iconic-systematicity (for a related discussion see Mompean, Fregier, and Valenzuela 2020).

4 | Open Questions

4.1 | Effects on Language Processing

Researchers have begun to examine whether iconic words are processed differently than non-iconic words. This could serve to demonstrate the psychological reality of iconicity. There has been evidence that onomatopoeia are processed faster than matched non-onomatopoeia (Meteyard et al. 2015; Sidhu, Vigliocco, and Pexman 2020; cf. Peeters 2016; Vigliocco et al. 2020). The limited evidence regarding crossmodal iconicity has been equivocal. Aryani and Jacobs (2018) conducted a semantic decision task in which participants categorised German words as exciting or calming. They found that participants were faster and more accurate when words were iconic (e.g., an exciting meaning and containing exciting sounds) compared to non-iconic. In this study, non-iconic words were specifically *anti*-iconic in that their sounds were related to the opposite of their meanings (e.g., an exciting meaning with a calming sound). Sidhu and Pexman (2022) used a similar paradigm to investigate size iconicity. Participants categorised nouns as large or small, and words contained sounds associated with either largeness or smallness. However, the authors found no evidence that iconic words were processed faster. There is a need for studies investigating effects of other kinds of iconic-systematicity.

It is worth pausing to consider the mechanism by which crossmodal iconicity could affect word processing. Sidhu, Vigliocco, and Pexman (2020) speculated that in a word processing system including orthographic, phonological and semantic units, iconic words might have special links between phonological and semantic units. This could be because in iconic words, phonological units can be connected to semantic units more directly (i.e., requiring less mediation to associate different kinds of information). Meteyard et al. (2015) proposed that iconicity could have an effect via links from phonological units to modality-specific sensorimotor features comprising a word's meaning, in a distributed network of language processing. Both explanations could be applied to crossmodal iconicity. Finally, it is notable that studies of crossmodal iconicity have used semantic decision tasks with the relevant dimensions (e.g., "exciting" and "calming") as responses. This allows for the possibility that these effects reflect an impact on decision making/response selection, rather than semantic processing per se.

A final consideration is how *systematicity* might affect processing in iconic-systematicity. There is evidence that systematicity in the absence of iconicity can affect language processing. Farmer, Christiansen, and Monaghan (2006) found that participants were faster to process sentences when target words were typical of their part of speech (i.e., verbs or nouns). Similarly, Reilly et al. (2012) found that participants were faster to categorise words as concrete or abstract when their forms were typical of their category. The authors suggested that in word form can serve as a cue to meaning prior to lexical access. Note that this is distinct from the previously reviewed theories of iconicity's effect on processing, in which lexical access is facilitated. Future studies of iconic-systematicity on language processing will need to disentangle the effects of these two properties.

4.2 | Causal Directionality

Studies of iconic-systematicity that predict participant ratings (e.g., of size, or shape) face a theoretical problem. Thus far we have assumed that phoneme associations have affected the

structure of the lexicon. However, another possibility is that these associations affected ratings of the relevant dimension. For instance, Sidhu et al. (2021) found that the phoneme /t/ was more common in words for spikier objects. Their interpretation was that the iconic link between /t/ and spikiness led to it becoming more common in words for spiky things. However, an alternate interpretation could be that the presence of the phoneme /t/ caused participants to rate an object as being spikier than they otherwise would have. There does seem to be some evidence of this. For instance, "tarantula" was judged to be spikier than "spider" in their dataset, despite the two being similarly shaped. This may have been because of the sounds the two words contain. Of course there are many counter examples, and this interpretation cannot explain the effect entirely. In fact, Sidhu et al. demonstrated that the association between word sound and shape was no stronger for more ambiguously shaped objects. One would have expected shape ratings for these items to be especially vulnerable to their sound. Nevertheless, future studies should be on guard for the possibility that phoneme associations can affect human ratings serving as dependent variables. They might supplement participant ratings with more objective measures of the relevant dimension.

Another consideration is that pseudoword ratings might be affected by patterns in language. Earlier I stated that pseudoword ratings serve to isolate the crossmodal associations of a given phoneme and bolster the case that a given pattern is iconic. For instance, the fact that pseudowords containing /t/ are rated spikier was used as justification by Sidhu et al. (2021) that the systematic pattern they found (i.e., /t/ being more common in words for spiky objects) was iconic. However, that systematic pattern could have contributed to pseudoword ratings, resulting in a circular argument. Akita and Imai (2022) refer to perceptions of iconicity originating from patterns in the lexicon as "secondary iconicity". Pseudoword rating studies done across cultures, or at early ages, could help provide evidence that phoneme associations are due primarily to crossmodal analogies rather than language patterns.[5]

4.3 | Emergence of Systematic Iconicity

A crucial question for the field is how iconic-systematicity emerges. Some have suggested that iconicity played a role in the earliest stages of language evolution (for a review, see Imai and Kita 2014). According to this theory, the resemblance between form and meaning was foundational for the emergence of spoken language. One could propose that the existence of iconic-systematicity is a vestige of such a proto-language. There is laboratory evidence that iconicity can affect the creation of novel words. Vinson et al. (2021) demonstrated that when asked to invent a word for a round or spiky shape, participants' responses tended to be iconic. There is also evidence that when individuals are asked to communicate a meaning using non-linguistic vocalizations, they make use of iconicity (Perlman, Dale, and Lupyan 2015; Perlman and Cain 2014), and that this helps naïve listener's interpretations (Perlman and Lupyan 2018).

Of course, the earliest stages of language emergence are beyond direct empirical study. So, while remaining agnostic to the origins of words, we might look to how words change over time. If iconicity conveys some benefit to language users, there may be a bias for a lexicon to gradually become more iconic. I have already reviewed evidence that the processing of iconic words may be facilitated. In addition, Imai and Kita (2014) argue that iconicity helps language learners form a link between sound and meaning, thereby facilitating acquisition of iconic words. Indeed, the earliest acquired words tend to be high in iconicity (Perry, Perlman, and Lupyan 2015; Sidhu et al. 2022; for a review of this topic, see A. K. Nielsen and Dingemanse 2021). Studies have also shown that adults have an easier time remembering the meanings of newly learnt foreign words if they are iconic (Lockwood, Dingemanse, and Hagoort 2016). All of these factors could provide a survival advantage for iconic words.

There is indeed evidence that is consistent with the argument that language becomes more iconic over time. Researchers have used iterated learning experiments to study language evolution in the lab (for a review see Kirby, Griffiths, and Smith 2014). A given participant is taught a set of words and definitions, and their later recall becomes the learning set for the next participant, and so on. Because participants' recall is not perfect, the "language" changes over time. This is argued to be an idealised version of how language changes as it is transmitted. Several iterated learning studies have indeed shown that iconic-systematicity emerges over time (Erben Johansson, Carr, and Kirby 2021; Vinson et al. 2021).

Instead of forms changing, another possibility is that meanings could change over time, affected by their forms. This is a possibility that has received less attention. Bolinger (1950; as cited by Nuckolls 1999) suggested that the word "pick" developed its sense of "picking on someone" due to the attraction of its neighours "peck" and "poke". Recently, Haslett and Cai (2022; see also 2023) demonstrated that perception of a pseudoword's meaning is influenced by its real word orthographic neighbours. This could be a mechanism by which iconic-systematicity reinforces itself, once already present. A related possibility is a word's form itself (as opposed to its neighbours) could lead to changes in its meaning. However, I am not aware of any work on this possibility.

Once iconic forms emerge in language, researchers have speculated that they may be resistant to change (for a review, see Nuckolls 1999). For example, Mithun (1982) discusses instances in which imitative words in Iroquoian languages did not undergo sound changes that would have been expected. However, this notion has been called into question by a recent survey of English which found that the majority of iconic words in English have changed their forms over time (see Flaksman 2017). Finally, as already mentioned, Winter et al. (2022) found that the pattern in which /r/ appears in words for roughness was present in Proto-Indo European roots from at least 6000 years ago, suggesting a perseverance of iconicity.

It will be important for future work to explore the emergence, and maintenance of iconic-systematicity in particular, in language. An ideal way to examine this would be to investigate snapshots of language at different points in history (see Flaksman 2017).

4.4 | Generalisation to Other Languages

Several of the studies reviewed have involved languages beyond English. Winter et al. (2022) explored the /r/-roughness link across 332 languages, Yu, McBeath, and Glenberg (2021) examined the gleam-glum effect in Mandarin speakers, and Aryani et al. (2018) focused on German (for examples of iconicity in translations of a particular concept across languages, see Blasi et al. 2016; N. Johansson, Anikin, and Aseyev 2020; Joo 2020). Nevertheless, there is a great deal of work to be done examining iconic-systematicity in other languages. Importantly, we would not expect the same patterns to emerge across languages. Because iconicity is in the eye of the beholder, linguistic and cultural factors shape perceptions of both form and meaning, and thus perceptions of iconic relations among them (see Occhino et al. 2017). While some iconic relationships may be so salient as to emerge with regularity across cultures (Ćwiek et al. 2022), this need not always be the case. Structural differences between languages could also affect the emergence of iconicity. For instance, Perry, Perlman, and Lupyan (2015) found differences between English and Spanish in terms of how iconicity manifested in different parts of speech. This was attributed to how either language conveyed manner of movement (i.e., within the verb itself or in a prepositional phrase). Thus, it will be important for future work to examine both generalizations and variations in these effects.

5 | Conclusion

Researchers have found that iconic-systematicity exists for the dimensions of shape, size, texture, valence and arousal. The field has answered the call from Sapir nearly 100 years ago and discovered that iconicity is a measurable undercurrent in the lexicon. Though, it is one of many, and its appearance can be sporadic. We are only beginning to understand the origin, extent, and effects of this phenomenon.

Acknowledgements

The author thanks Emiko Muraki for her very helpful feedback on an earlier draft.

Endnotes

[1] Note that while spoken language is the focus here, iconicity exists beyond spoken languages. For example, there is a vast literature on sound symbolism in sign language (see Emmorey 2014).

[2] However, it is worth pointing the reader to the fascinating work that has been done across languages (for a review, see N. Johansson et al. 2020). A study by Blasi et al. (2016) looked for patterns in the phonemes contained in translations of 100 basic vocabulary words across more than 6000 languages. Indeed, there was some evidence of crossmodal iconicity. For example, words for "small" tended to contain the phoneme /i/. This cross-language approach has found patterns in the words for body parts, colour, and spatial deixis (see N. Johansson, Anikin, and Aseyev 2020; Joo 2020).

[3] Though, the version of this paper that was accessed was from 1933, containing additions after the publications of Sapir (1929).

[4] While I focus here on their prediction of arousal, they did find that words with a more positive valence tended to have a lower F1 frequency, as well as an overall lower and less varied spectral center of gravity.

[5] Recall that iconicity is considered to exist in the eye of the beholder, as a perceived resemblance between construals of form and meaning (see Occhino et al. 2017). Crucially, these construals will be shaped by linguistic and cultural experience. If this perspective is applied to pseudoword ratings, then it may be neither possible nor meaningful to disentangle the contributions of crossmodal analogies and language patterns. Nevertheless, the suggestions above may provide a measure of phoneme associations that give less weight to language patterns.

References

Adelman, J. S., Z. Estes, and M. Cossu. 2018. "Emotional Sound Symbolism: Languages Rapidly Signal Valence via Phonemes." *Cognition* 175: 122–130. https://doi.org/10.1016/j.cognition.2018.02.007.

Akita, K., and M. Imai. 2022. "The Iconicity Ring Model for Sound Symbolism." *Iconicity in Language and Literature*: 27–46. https://doi.org/10.1075/ill.18.02aki.

Aryani, A., M. Conrad, D. Schmidtke, and A. Jacobs. 2018. "Why 'Piss' Is Ruder Than 'Pee'? the Role of Sound in Affective Meaning Making." *PLoS One* 13, no. 6: e0198430. https://doi.org/10.1371/journal.pone.0198430.

Aryani, A., and A. Jacobs. 2018. "Affective Congruence Between Sound and Meaning of Words Facilitates Semantic Decision." *Behavioral Sciences* 8, no. 6: 56. https://doi.org/10.3390/bs8060056.

Bergen, B. K. 2004. "The Psychological Reality of Phonaesthemes." *Language* 80, no. 2: 290–311. https://doi.org/10.1353/lan.2004.0056.

Berlin, B. 1992. *Ethnobiological Classification: Principles of Categorization of Plants and Animals in Traditional Societies*. Princeton Univ. Press.

Blasi, D. E., S. Wichmann, H. Hammarström, P. F. Stadler, and M. H. Christiansen. 2016. "Sound–Meaning Association Biases Evidenced Across Thousands of Languages." *Proceedings of the National Academy of Sciences* 113, no. 39: 10818–10823. https://doi.org/10.1073/pnas.1605782113.

Cuskley, C. 2013. "Mappings Between Linguistic Sound and Motion." *Public Journal of Semiotics* 5, no. 1: 39–62: Article 1. https://doi.org/10.37693/pjos.2013.5.9651.

Ćwiek, A., R. Anselme, D. Dediu, et al. 2024. "The Alveolar Trill Is Perceived as Jagged/Rough by Speakers of Different Languages."*The Journal of the Acoustical Society of America* 156, no. 5: 3468–3479. https://doi.org/10.1121/10.0034416.

Ćwiek, A., S. Fuchs, C. Draxler, et al. 2022. "The Bouba/kiki Effect Is Robust Across Cultures and Writing Systems." *Philosophical Transactions of the Royal Society B: Biological Sciences* 377, no. 1841: 20200390. https://doi.org/10.1098/rstb.2020.0390.

Dingemanse, M., D. E. Blasi, G. Lupyan, M. H. Christiansen, and P. Monaghan. 2015. "Arbitrariness, Iconicity, and Systematicity in Language." *Trends in Cognitive Sciences* 19, no. 10: 603–615. https://doi.org/10.1016/j.tics.2015.07.013.

Dingemanse, M., M. Perlman, and P. Perniss. 2020. "Construals of Iconicity: Experimental Approaches to Form–Meaning Resemblances in Language." *Language and Cognition* 12, no. 1: 1–14. https://doi.org/10.1017/langcog.2019.48.

D'Onofrio, A. 2014. "Phonetic Detail and Dimensionality in Sound-Shape Correspondences: Refining the Bouba-Kiki Paradigm." *Language and Speech* 57, no. 3: 367–393. https://doi.org/10.1177/0023830913507694.

Emmorey, K. 2014. "Iconicity as Structure Mapping." *Philosophical Transactions of the Royal Society B: Biological Sciences* 369, no. 1651: 20130301. https://doi.org/10.1098/rstb.2013.0301.

Erben Johansson, N., J. W. Carr, and S. Kirby. 2021. "Cultural Evolution Leads to Vocal Iconicity in an Experimental Iterated Learning Task." *Journal of Language Evolution* 6, no. 1: 1–25. https://doi.org/10.1093/jole/lzab001.

Farmer, T. A., M. H. Christiansen, and P. Monaghan. 2006. "Phonological Typicality Influences On-Line Sentence Comprehension." *Proceedings of the National Academy of Sciences* 103, no. 32: 12203–12208. https://doi.org/10.1073/pnas.0602173103.

Flaksman, M. (2017). Iconic Treadmill Hypothesis: The Reasons Behind Continuous Onomatopoeic Coinage, A. Zirker, M. Bauer, O. Fischer, & C. Ljungberg (Eds.), *Iconicity in Language and Literature* (Vol. 15). John Benjamins Publishing Company. https://doi.org/10.1075/ill.15.02fla

Fort, M., I. Lammertink, S. Peperkamp, A. Guevara-Rukoz, P. Fikkert, and S. Tsuji. 2018. "Symbouki: A Meta-Analysis on the Emergence of Sound Symbolism in Early Language Acquisition." *Developmental Science* 21, no. 5: e12659. https://doi.org/10.1111/desc.12659.

Gasser, M. 2004. "The Origins of Arbitrariness in Language." *Proceedings of the Annual Meeting of the Cognitive Science Society* 26.

Haslett, D. A., and Z. G. Cai. 2022. "New Neighbours Make Bad Fences: Form-Based Semantic Shifts in Word Learning." *Psychonomic Bulletin & Review* 29, no. 3: 1017–1025. https://doi.org/10.3758/s13423-021-02037-1.

Haslett, D. A., and Z. G. Cai. 2023. "Similar-Sounding Words Flesh Out Fuzzy Meanings." *Journal of Experimental Psychology: General* 152, no. 8: 2359–2368. https://doi.org/10.1037/xge0001409.

Householder, F. W. 1946. "On the Problem of Sound and Meaning, and English Phonesthemes." *Word* 2: 83–84.

Imai, M., and S. Kita. 2014. "The Sound Symbolism Bootstrapping Hypothesis for Language Acquisition and Language Evolution." *Philosophical Transactions of the Royal Society B: Biological Sciences* 369, no. 1651: 20130298. https://doi.org/10.1098/rstb.2013.0298.

Jespersen, O. 1922. "The Symbolic Value of the Vowel I." *Philologica* 1: 1–19.

Johansson, N., A. Anikin, and N. Aseyev. 2020. "Color Sound Symbolism in Natural Languages." *Language and Cognition* 12, no. 1: 56–83. https://doi.org/10.1017/langcog.2019.35.

Johansson, N. E., A. Anikin, G. Carling, and A. Holmer. 2020. "The Typology of Sound Symbolism: Defining Macro-Concepts via Their Semantic and Phonetic Features." *Linguistic Typology* 24, no. 2: 253–310. https://doi.org/10.1515/lingty-2020-2034.

Johnson, R. C. 1967. "Magnitude Symbolism of English Words." *Journal of Verbal Learning and Verbal Behavior* 6, no. 4: 508–511. https://doi.org/10.1016/S0022-5371(67)80008-2.

Joo, I. 2020. "Phonosemantic Biases Found in Leipzig-Jakarta Lists of 66 Languages." *Linguistic Typology* 24, no. 1: 1–12. https://doi.org/10.1515/lingty-2019-0030.

Katz, A. N. 1986. "Meaning Conveyed by Vowels: Some Reanalyses of Word Norm Data." *Bulletin of the Psychonomic Society* 24, no. 1: 15–17. https://doi.org/10.3758/BF03330490.

Kirby, S., T. Griffiths, and K. Smith. 2014. "Iterated Learning and the Evolution of Language." *Current Opinion in Neurobiology* 28: 108–114. https://doi.org/10.1016/j.conb.2014.07.014.

Knoeferle, K., J. Li, E. Maggioni, and C. Spence. 2017. "What Drives Sound Symbolism? Different Acoustic Cues Underlie Sound-Size and Sound-Shape Mappings." *Scientific Reports* 7, no. 1: 5562: Article 1. https://doi.org/10.1038/s41598-017-05965-y.

Köhler, W. 1929. *Gestalt Psychology*. Liverlight.

Lockwood, G., and M. Dingemanse. 2015. "Iconicity in the Lab: A Review of Behavioral, Developmental, and Neuroimaging Research into Sound-Symbolism." *Frontiers in Psychology* 6. https://doi.org/10.3389/fpsyg.2015.01246.

Lockwood, G., M. Dingemanse, and P. Hagoort. 2016. "Sound-Symbolism Boosts Novel Word Learning." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 42, no. 8: 1274–1281. https://doi.org/10.1037/xlm0000235.

Mallory, J. P., and D. Q. Adams. 2006. *The Oxford Introduction to Proto-Indo-European and the Proto-Indo-European World*. USA: Oxford University Press.

Meteyard, L., E. Stoppard, D. Snudden, S. F. Cappa, and G. Vigliocco. 2015. "When Semantics Aids Phonology: A Processing Advantage for Iconic Word Forms in Aphasia." *Neuropsychologia* 76: 264–275. https://doi.org/10.1016/j.neuropsychologia.2015.01.042.

Mithun, M. 1982. "The Synchronic and Diachronic Behavior of Plops, Squeaks, Croaks, Sighs, and Moans." *International Journal of American Linguistics* 48, no. 1: 49–58. https://doi.org/10.1086/465712.

Mompean, J. A., A. Fregier, and J. Valenzuela. 2020. "Iconicity and Systematicity in Phonaesthemes: A Cross-Linguistic Study." *Cognitive Linguistics* 31, no. 3: 515–548. https://doi.org/10.1515/cog-2018-0079.

Monaghan, P., K. Mattock, and P. Walker. 2012. "The Role of Sound Symbolism in Language Learning." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 38, no. 5: 1152–1164. https://doi.org/10.1037/a0027747.

Newman, S. S. 1933. "Further Experiments in Phonetic Symbolism." *American Journal of Psychology* 45, no. 1: 53–75. https://doi.org/10.2307/1414186.

Nielsen, A., and D. Rendall. 2011. "The Sound of Round: Evaluating the Sound-Symbolic Role of Consonants in the Classic Takete-Maluma Phenomenon." *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale* 65, no. 2: 115–124. https://doi.org/10.1037/a0022268.

Nielsen, A. K., and M. Dingemanse. 2021. "Iconicity in Word Learning andBeyond: A Critical Review." *Language and Speech* 64, no. 1: 52–72. https://doi.org/10.1177/0023830920914339.

Nuckolls, J. B. 1999. "The Case for Sound Symbolism." *Annual Review of Anthropology* 28, no. 1: 225–252. https://doi.org/10.1146/annurev.anthro.28.1.225.

Occhino, C., B. Anible, E. Wilkinson, and J. P. Morford. 2017. "Iconicity Is in the Eye of the Beholder: How Language Experience Affects Perceived Iconicity." *Gesture* 16, no. 1: 100–126. https://doi.org/10.1075/gest.16.1.04occ.

Peeters, D. 2016. Processing Consequences of Onomatopoeic Iconicity in Spoken Language Comprehension. https://pure.mpg.de/pubman/faces/ViewItemOverviewPage.jsp?itemId=item_2304565.

Perlman, M., and A. A. Cain. 2014. "Iconicity in Vocalization, Comparisons With Gesture, and Implications for Theories on the Evolution of Language." *Gesture* 14, no. 3: 320–350. https://doi.org/10.1075/gest.14.3.03per.

Perlman, M., R. Dale, and G. Lupyan. 2015. "Iconicity Can Ground the Creation of Vocal Symbols." *Royal Society Open Science* 2, no. 8: 150152. https://doi.org/10.1098/rsos.150152.

Perlman, M., and G. Lupyan. 2018. People Can Create Iconic Vocalizations to Communicate Various Meanings to Naïve Listeners Scientific Reports. https://www.nature.com/articles/s41598-018-20961-6.

Perniss, P., R. L. Thompson, and G. Vigliocco. 2010. "Iconicity as a General Property of Language: Evidence From Spoken and Signed Languages." *Frontiers in Psychology* 1. https://doi.org/10.3389/fpsyg.2010.00227.

Perry, L. K., M. Perlman, and G. Lupyan. 2015. "Iconicity in English and Spanish and its Relation to Lexical Category and Age of Acquisition."

PLoS One 10, no. 9: e0137147. https://doi.org/10.1371/journal.pone.0137147.

Ramachandran, V. S., and E. M. Hubbard. 2001. "Synaesthesia—A Window Into Perception, Thought and Language." *Journal of Consciousness Studies* 8, no. 12: 3–34.

Reilly, J., C. Westbury, J. Kean, and J. E. Peelle. 2012. "Arbitrary Symbolism in Natural Language Revisited: When Word Forms Carry Meaning." *PLoS One* 7, no. 8: e42286. https://doi.org/10.1371/journal.pone.0042286.

Sapir, E. 1929. "A Study in Phonetic Symbolism." *Journal of Experimental Psychology* 12, no. 3: 225–239. https://doi.org/10.1037/h0070931.

Sidhu, D. M., K. Deschamps, J. S. Bourdage, and P. M. Pexman. 2019. "Does the Name Say it All? Investigating Phoneme-Personality Sound Symbolism in First Names." *Journal of Experimental Psychology: General* 148, no. 9: 1595–1614. https://doi.org/10.1037/xge0000662.

Sidhu, D. M., and P. M. Pexman. 2017. "A Prime Example of the Maluma/Takete Effect? Testing for Sound Symbolic Priming." *Cognitive Science* 41, no. 7: 1958–1987. https://doi.org/10.1111/cogs.12438.

Sidhu, D. M., and P. M. Pexman. 2018. "Five Mechanisms of Sound Symbolic Association." *Psychonomic Bulletin & Review* 25, no. 5: 1619–1643. https://doi.org/10.3758/s13423-017-1361-1.

Sidhu, D. M., and P. M. Pexman. 2022. "Is a Boat Bigger Than a Ship? Null Results in the Investigation of Vowel Sound Symbolism on Size Judgements in Real Language." *Quarterly Journal of Experimental Psychology* 76, no. 1: 17470218221078299. https://doi.org/10.1177/17470218221078299.

Sidhu, D. M., G. Vigliocco, and P. M. Pexman. 2020. "Effects of Iconicity in Lexical Decision." *Language and Cognition* 12, no. 1: 164–181. https://doi.org/10.1017/langcog.2019.36.

Sidhu, D. M., G. Vigliocco, and P. M. Pexman. 2022a. "Higher Order Factors of Sound Symbolism." *Journal of Memory and Language* 125: 104323. https://doi.org/10.1016/j.jml.2022.104323.

Sidhu, D. M., C. Westbury, G. Hollis, and P. M. Pexman. 2021. "Sound Symbolism Shapes the English Language: The Maluma/takete Effect in English Nouns." *Psychonomic Bulletin & Review* 28, no. 4: 1390–1398. https://doi.org/10.3758/s13423-021-01883-3.

Sidhu, D. M., J. Williamson, V. Slavova, and P. M. Pexman. 2022b. "An Investigation of Iconic Language Development in Four Datasets." *Journal of Child Language* 49, no. 2: 382–396. https://doi.org/10.1017/S0305000921000040.

Strobl, C., A.-L. Boulesteix, A. Zeileis, and T. Hothorn. 2007. "Bias in Random Forest Variable Importance Measures: Illustrations, Sources and a Solution." *BMC Bioinformatics* 8, no. 1: 25. https://doi.org/10.1186/1471-2105-8-25.

Styles, S. J., and L. Gawne. 2017. "When Does Maluma/Takete Fail? Two Key Failures and a Meta-Analysis Suggest That Phonology and Phonotactics Matter." *I-Perception* 8, no. 4: 2041669517724807. https://doi.org/10.1177/2041669517724807.

Taylor, I. K., and M. M. Taylor. 1965. "Another Look at Phonetic Symbolism." *Psychological Bulletin* 64, no. 6: 413–427. https://doi.org/10.1037/h0022737.

Thompson, P. D., and Z. Estes. 2011. "Sound Symbolic Naming of Novel Objects Is a Graded Function." *Quarterly Journal of Experimental Psychology* 64, no. 12: 2392–2404. https://doi.org/10.1080/17470218.2011.605898.

Thorndike, E. L. 1945. "The Association of Certain Sounds With Pleasant and Unpleasant Meanings." *Psychological Review* 52, no. 3: 143–149. https://doi.org/10.1037/h0055510.

Trask, R. L., and L. Robert. 2000. *The Penguin Dictionary of English Grammar*. London: Penguin: (with Internet Archive). http://archive.org/details/penguindictionar0000unse_k2y7.

Vigliocco, G., Y. Zhang, N. Del Maschio, R. Todd, and J. Tuomainen. 2020. "Electrophysiological Signatures of English Onomatopoeia." *Language and Cognition* 12, no. 1: 15–35. https://doi.org/10.1017/langcog.2019.38.

Vinson, D., M. Jones, D. M. Sidhu, A. Lau-Zhu, J. Santiago, and G. Vigliocco. 2021. "Iconicity Emerges and Is Maintained in Spoken Language." *Journal of Experimental Psychology: General* 150, no. 11: 2293–2308. https://doi.org/10.1037/xge0001024.

Westbury, C., G. Hollis, D. M. Sidhu, and P. M. Pexman. 2018. "Weighing up the Evidence for Sound Symbolism: Distributional Properties Predict Cue Strength." *Journal of Memory and Language* 99: 122–150. https://doi.org/10.1016/j.jml.2017.09.006.

Winter, B., and M. Perlman. 2021. "Size Sound Symbolism in the English Lexicon." *Glossa: A Journal of General Linguistics* 6, no. 1: Article 1. https://doi.org/10.5334/gjgl.1646.

Winter, B., M. Perlman, L. K. Perry, and G. Lupyan. 2017. "Which Words Are Most Iconic?: Iconicity in English Sensory Words." *Interaction Studies* 18, no. 3: 443–464. https://doi.org/10.1075/is.18.3.07win.

Winter, B., M. Sóskuthy, M. Perlman, and M. Dingemanse. 2022. "Trilled/r/Is Associated With Roughness, Linking Sound and Touch Across Spoken Languages." *Scientific Reports* 12, no. 1: 1035: Article 1. https://doi.org/10.1038/s41598-021-04311-7.

Winter, B., G. Woodin, and M. Perlman. 2023. *Winter, Woodin & Perlman—Defining Iconicity for the Cognitive Sciences*. OSF. https://doi.org/10.31219/osf.io/5e3rc.

Yu, C. S.-P., M. K. McBeath, and A. M. Glenberg. 2021. "The Gleam-Glum Effect:/I:/versus/λ/Phonemes Generically Carry Emotional Valence." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 47, no. 7: 1173–1185. https://doi.org/10.1037/xlm0001017.