OXFORD

Cultural evolution leads to vocal iconicity in an experimental iterated learning task

Niklas Erben Johansson (iD)*,[1], **Jon W. Carr**[2] **and Simon Kirby**[3]

[1]Center for Language and Literature, Lund University, Helgonabacken 12, Lund, SE-223 62, Sweden, [2]Cognitive Neuroscience, International School for Advanced Studies, Via Bonomea, 265, 34136 Trieste, Italy and [3]Centre for Language Evolution, University of Edinburgh, Dugald Stewart Building, 3 Charles St, Edinburgh EH8 9AD, UK

*Corresponding author: niklas.erben_johansson@ling.lu.se, n.johansson@hotmail.com

Abstract

Experimental and cross-linguistic studies have shown that *vocal iconicity* is prevalent in words that carry meanings related to SIZE and SHAPE. Although these studies demonstrate the importance of vocal iconicity and reveal the cognitive biases underpinning it, there is less work demonstrating how these biases lead to the evolution of a sound symbolic lexicon in the first place. In this study, we show how words can be shaped by cognitive biases through cultural evolution. Using a simple experimental setup resembling the game *telephone*, we examined how a single word form changed as it was passed from one participant to the next by a process of *immediate iterated learning*. About 1,500 naïve participants were recruited online and divided into five condition groups. The participants in the CONTROL-group received no information about the meaning of the word they were about to hear, while the participants in the remaining four groups were informed that the word meant either BIG or SMALL (with the meaning being presented in text), or ROUND or POINTY (with the meaning being presented as a picture). The first participant in a transmission chain was presented with a phonetically diverse word and asked to repeat it. Thereafter, the recording of the repeated word was played for the next participant in the same chain. The sounds of the audio recordings were then transcribed and categorized according to six binary sound parameters. By modelling the proportion of vowels or consonants for each sound parameter, the SMALL-condition showed increases of FRONT UNROUNDED vowels and the POINTY-condition increases of ACUTE consonants. The results show that linguistic transmission is sufficient for vocal iconicity to emerge, which demonstrates the role non-arbitrary associations play in the evolution of language.

Key words: iconcitiy; sound symbolism; iterated learning; language evolution

1. Introduction

Languages have iconic structure—that is, some form of resemblance between sound and meaning—woven into the very core of the lexicon (Dingemanse et al. 2015; Blasi et al. 2016; Erben Johansson et al. 2020; Joo 2020). But how does such patterning enter languages and what explains its apparent universality? In this article, we use the experimental iterated learning paradigm to show how the cultural transmission of a single artificial word may converge on iconic sound-meaning

correspondences that closely reflect the kinds of patterns observed in natural languages.

1.1 Oppositional vocal iconicity

The number of studies on the genetically and areally independent, (near-)universal, non-arbitrary, and flexible associations between sounds and meanings has grown considerably in recent decades. This type of association is generally referred to as *vocal iconicity* or *motivated sound symbolism* (Cuskley and Kirby 2013). Several large cross-linguistic studies (Wichmann et al. 2010; Blasi et al. 2016; Erben Johansson et al. 2020), which in some cases incorporate data from thousands of languages, have identified a number of robust overrepresentations of sounds across languages in basic vocabulary items for concepts that are supposed to be more or less universal to all speakers of all languages (e.g. *tree*, *you*, *mother*, *eat*, *black*, *small*), both culturally and historically (Swadesh 1971; Goddard and Wierzbicka 2002). Collectively, these studies found iconic effects for a wide range of meanings consisting mostly of several fundamental nouns (e.g. ASHES, BREASTS, NOSE, and TONGUE) and verbs (e.g. TO BLOW, TO BITE, and TO SNEEZE), but also a few pronouns (I, WE, and YOU) and adjectives (RED, ROUND, and SMALL).

Experimental evidence has also covered a variety of meanings. For example, Maglio et al., (2014) found that front vowels, as opposed to back vowels, tend to be linked to conceptual precision in fictional city names. Fast speed has also been linked to front vowels and slow speed to back vowels in non-words when asking participants to describe the motion of a ball (Cuskley 2013). Anikin and Johansson (2018); Hamilton-Fletcher et al. (2018), Johansson et al. (2019) found a similar effect for tastes. Participants tend to assign lower F_1 and F_2 frequencies to salty taste samples and higher F_1 and F_2 frequencies to sweet taste samples. Additionally, there is a large body of studies, both experimental and cross-linguistic, on associations between different color parameters, such as lightness, saturation and hue, and acoustic parameters, such as pitch, energy spectrum, vowel formants, and loudness, in great apes, infants, toddlers, adults, synesthetes, non-synesthetes, etc. (Anikin and Johansson 2018; Hamilton-Fletcher et al. 2018; Johansson et al. 2019).

Evidently, vocal iconicity seems to be prevalent in the core of the lexicon; however, with the exception of color, most experimental studies on vocal iconicity have been focused on the SIZE- and SHAPE-dimensions, with the sound side usually conveyed through made-up nonwords. Almost one hundred years ago, Sapir (1929) conducted a study on size-based vocal iconicity which showed that 80% of almost 500 participants preferred to associate a small table with the phonetic form /mil/ and a large table with the form /mal/. Similarly, Köhler (1929) investigated shape-based vocal iconicity by asking participants to match a round, amoeba-like shape and a pointy, star-like shape with either /takete/ or /baluma/ (later replaced by /maluma/ in his 1947 study). Most of the participants thought that the best fit for the round shape was the word containing voiced sounds and the pointy shape was accordingly paired with the word containing unvoiced sounds. Köhler's (1929) work was later built on by several scholars (e.g. Rogers and Ross 1975; O'Boyle and Tarte 1980; Lindauer 1990; Holland and Wertheimer 2016; Bross 2018; for a review see Lockwood and Dingemanse (2015), and associations between round shapes and phonetic forms, such as /maluma/ or /bouba/, and associations between pointy shapes and phonetic forms, such as /takete/ or /kiki/, have since be found to hold for around 90% of participants with a wide range of first languages (Styles and Gawne 2017).

Similar studies on words from natural languages have demonstrated iconic effects in a wide range of semantically opposite meanings. Newman (1933) found a correspondence between the articulation and acoustics of vowels and those vowels' perceived size and brightness. This showed that vowels pronounced at the back of the mouth had a lower acoustic frequency, which was also judged to be larger and darker. Johnson (1967) showed that when participants were tasked with coming up with English words to denote small and large size, the vowel quality in the words correlated with the words meanings', ranging from smallest /i/, followed by /e/, /a/ and /u/, to largest /o/. This study was later expanded to include Mandarin Chinese and Thai which also yielded similar results (Huang et al. 1969). Fónagy (1963) compared /i/ and /u/ in Hungarian and concluded that /i/ was considered quicker, smaller, prettier, friendlier, and harder than /u/, while /u/ was perceived as thicker, hollower, darker, sadder, blunter, more bitter, and stronger than /i/ (in both children and adults). Taylor and Taylor (1962) and Taylor (1963) found iconic effects for big-small, active-passive, warm-cold, and pleasant-unpleasant in four unrelated languages, and Gebels (1969) found that speakers of five different languages could correctly match the meaning of sensory words from the other languages above chance level.

Perhaps, the most widely known type of vocal iconicity is onomatopoeia (i.e. human imitations of real-world sounds with varying similarity to the source sound), which has been referred to as *imagic*, *absolute* or *imitative iconicity* (Hinton et al. 1994; Dingemanse

2011; Carling and Johansson 2014; D'Onofrio 2014; Dingemanse et al. 2015). For example, the English word *cuckoo* is a direct imitation of the calls produced by the cuckoo but produced through the filter of the human vocal apparatus and linguistic sound system. However, in contrast to onomatopoeia, the type of vocal iconicity usually investigated experimentally involves referents that are based on senses other than hearing, such as size, shape, deixis, or color, and can in most cases be classified as *relative* or *word-relational diagrammatic iconicity*. Relative iconicity is constructed by mapping semantic contrasts to phonetic contrasts which are somehow similar to each other. This usually includes binary semantic meanings that can easily be placed in opposition to each other (FAST-SLOW, BIG-SMALL, ROUND-POINTY, etc.) and phonetic attributes that can be perceived to belong to a gradable scale (e.g. voicing, quality, quantity, tone, volume, etc.). For example, if SMALL is mapped to front unrounded vowels and BIG is mapped to back rounded vowels, these parallel sound-meaning associations add relations between the semantic and phonetic parameters to the internal relations within the semantic parameter SIZE (between BIG and SMALL) and the phonetic parameter roundedness (between unrounded and rounded vowels). Ohala (1994) argues that the so-called *frequency code* (see also Rendall et al. 2005) could be the underlying mechanism responsible for associations of this type. It states that since the size of the resonance chamber of an animal dictates the fundamental frequency of that animal's vocalizations, the sounds that the animal produces can be utilized in various ways to evoke properties such as size. This works according to the same principle as erecting feathers or fur in threatening situations to seem larger or cowering when wanting to submit. Ohala therefore argues that most animals, and maybe humans, perceive low and/or falling fundamental frequencies of vocalizations such as growling as large, authoritative, confident, dominant, or distant, and high and/or rising fundamental frequencies of vocalizations, such as whining as small, polite, questioning, dependent, or near.

1.2 Confounds in vocal iconicity experiments

Based on these studies, sounds associated with meanings belonging to the SIZE and SHAPE-domains, which are the focus of the current study, can be summed up into a few general groups. The meaning SMALL has been reliably associated with voiceless consonants and vowels with a low first formant or high second formant (e.g. [i] or [a]) while the meaning BIG has been associated with voiced consonants and vowels with a high first formant or low

second formant (e.g. [u] or [ɑ]). The associations with SHAPE meanings correlate partially to those found for SMALL and BIG, but there are some crucial differences. While voiceless obstruents (e.g. [t] and [k]) and vowels with a high second formant (e.g. [i] and [e]) have been found to be associated with the meaning POINTY, the meaning ROUND is associated with voiced consonants (e.g. [b]), sonorants (e.g. [m] and [l]), and vowels with a low second formant (e.g. [u]).

There are, however, a number of confounding factors present in most of the previous studies on SIZE- and SHAPE-symbolism. In order to correspond to the oppositional semantic relationships constituted by the SMALL-BIG and POINTY-ROUND pairs, previous studies have included constructed non-word stimuli which consisted of contrasting sounds (e.g. [m] vs. [k] or [u] vs. [i]). The selected sounds have, however, generally conflated several phonetic parameters which make it harder to distinguish which factors are driving the sound-meaning associations. For example, the association between [u] and round shapes might primarily be attributed to the roundedness of the vowel, but it is possible that other parameters, e.g. vowel height and backness, play a role as well. Furthermore, despite the large number of studies that have found supporting evidence for the bouba-kiki effect, there are two reported cases where the effect has failed (Rogers and Ross 1975; Styles and Gawne 2017). Both of these were, however, conducted with participants speaking languages in which the stimuli words were potentially not phonologically possible. This could have led to issues with parsing the linguistic strings, which, in turn, could also result in a breakdown in the mapping between sounds and meanings. This raises questions about the strength of the bouba-kiki effect as well as the influence of language-specific phonological makeup. Thus, while binary distinctions between sounds ought to be useful for studying binary semantic pairs, they should be broad enough to accommodate speakers with different first languages, as well as capture individual parameters which are also articulatory, acoustically, and iconically relevant.

Vowels can be primarily divided according to height, backness and roundedness, which also loosely correspond to the first three formants, F_1, F_2, and F_3, and thus, cover most of the variation used for distinguishing vowel segments across languages (Ladefoged 2001: 32–36). Furthermore, energy level differences in F_1 and F_2 have been iconically linked to size, distance, dominance, etc., while the roundedness of F_3 has been linked to shape. Consonants are considerably more articulatorily diverse than vowels, but can generally be divided according to voicing, manner of articulation and

position of articulation. The distinction between voiceless and voiced sounds is self-explanatory and cuts through all consonants and is used phonemically in most languages(Ladefoged 2001: 63–65; Ladefoged and Maddieson 1996: 44–46). In addition, it is, like F_1 and F_2, iconically associated with a number of meanings, such as size (Ohala 1994). Manner of articulation include a wide variety of sound types, but one of the most fundamental ways of classifying consonants is to distinguish between sonorants and obstruents. This distinction has also been employed for several iconicity experiments because the contrast between sonorants' continuous, non-turbulent airflow, and obstruents' obstructed airflow could iconically evoke, for example, noisiness versus smoothness or other related meanings. Position of articulation also includes a range of different sounds which can be difficult to fit in a binary distinction without basing the distinction on a specific marked feature and grouping the remaining features in a contrastive group. However, the distinction between grave and acute sounds (Jakobson et al. 1953) differentiates between perceptually sharper and duller sounds which has also been linked to iconic associations (Lapolla 1994). Grave sounds include consonants produced by using soft tissue secondary articulators, notably the lips and the area from the soft palate and back, while acute sounds include consonants produced using the hard palate as a secondary articulator.

Related to the binary division of sounds, the inherent structure of relative diagrammatic iconicity in which two poles of phonetic and semantic parameters are mapped in parallel have led researchers to create stimuli consisting of premade non-words. However, this methodological setup does not explore which sounds are actually relevant for identifying iconic mappings and might in some cases yield incorrect information about the strength of these iconic effects. Nielsen and Rendall (2012) conducted a learning experiment in which one group of participants was taught to combine non-words with iconically congruent figures and one group was taught to combine non-words with iconically incongruent figures, after which all participants were subjected to random single word-figure combinations and asked to judge the combinations as "correct" or not. The results revealed that while the participants in the incongruent group performed at chance level, the congruent group performed only modestly (53.3% correct). This therefore suggests that the iconic bias might be weaker than demonstrated by previous studies and that the forced choice paradigm could inflate weak effects (Dingemanse et al. 2015).

Furthermore, premade non-words contain both vowels and consonants, but studies have shown that the effects of vowels and consonants might differ in iconic strength. Ahlner and Zlatev (2010) investigated the bouba-kiki effect by selecting vowels and consonants that had been reported to contrast iconically and then created four sets of non-word types. Two of these types were iconically congruent (e.g. [titi] for POINTY and [mumu] for ROUND), while the other word types were iconically incongruent (e.g. e.g. [tutu] for POINTY and [mimi] for ROUND). Participants were then asked to match these words to a pointy or round shape, virtually identical to those used by Ramachandran and Hubbard (2001). The results showed a clear preference for the iconically congruent words, and this effect was stronger in the words with congruent consonants, which might indicate that consonants play a more important role in this iconic mapping. Similarly, Nielsen and Rendall (2013) demonstrated a stronger preference for plosives and unrounded vowels in pointy figures, as opposed to sonorants and rounded vowels for the round figures.

Finally, another issue in this type of experiment is orthography, since the vast majority of tested participants are from literate societies. This could result in the shapes of letters in text stimuli having an effect on responses. For example, Nielsen and Rendall (2011) investigated the bouba-kiki effect using stimuli non-words conveyed through text in which they found iconic effects for consonants but not vowels. The experiment was then rerun using auditory stimuli through a text-to-speech synthesizer to exclude any orthographic bias, and they found the same general pattern regarding consonants and vowels, although the overall effect was weaker than the text stimuli. Furthermore, Cuskley et al. (2017) found that orthography seems to be a major confounding factor for associations between sounds and shapes—the voiced/voiceless distinction, for example, which they argue had not been sufficiently controlled for in most previous studies. By both testing how well literate participants matched abstract shapes to non-words in written form along with spoken representations, and how well they matched the shapes to purely auditory non-words, Cuskley et al. showed that the curvature of letters can significantly influence the perceived roundedness of shapes in sound–shape associations.

However, Hamilton-Fletcher et al. (2018) showed that these types of correspondences might be more complex. While pitch-shape correspondences required visual

experience to emerge in blind participants, pitch-size and pitch-weight were found to be unaffected by visual experience, and pitch-texture and pitch-softness even seemed to emerge or grow stronger with blindness. Thus, visual experience cannot solely explain why people with limited multisensory interactions have multimodal perception. Instead, this could be attributed to other factors such as neuroplasticity.

1.3 Vocal iconicity through iterated learning

Some general conclusions can be drawn from the different approaches that the studies we have reviewed have employed. In the bouba-kiki effect, both vowels and consonants seem to play a role, which illustrates the value of thoroughly investigating how different sounds are mapped to different meanings. Furthermore, previous studies, with some notable exceptions (e.g. Jones et al. 2014; Tamariz et al. 2017 , described below), have typically relied on experimental paradigms in which participants are asked to associate meanings with a set of non-words or syllables that are predefined. This means that while the bouba-kiki effect seems to be more or less universal, it is also subjective in nature, given that each individual participant is asked to combine meanings with sounds that may or may not adequately fit his or her intuition or phonology. We therefore wanted to investigate the cognitive biases that lie at the core of vocal iconicity by using a methodological approach that focuses on the transmission of vocal iconicity through the language filters of participants with a wide range of native languages, but which also excludes orthographic influence as much as possible. This approach would then allow us to get a more holistic picture of the bouba-kiki effect by revealing differences to the results of previous studies.

One way of achieving this is to use methods that are designed to study how systems, for example languages, change over time, such as the *iterated learning paradigm* (Kirby 2001; Kirby and Hurford 2002; Simon Kirby et al. 2008, 2015). In iterated learning studies, some form of information, such as words, music, or drawings, is transmitted from one participant to another, with the learner at generation *i* producing behavior that is input to the learner at generation *i* + 1. Together, several generations of such learners form a 'transmission chain'. At its core, the iterated learning paradigm is reliant on the fact that information tends to be lost during the transmission process (Spike et al. 2017), causing the object of study to change in ways that reflect the learner's cognitive biases, whatever those biases happen to be, and the dynamics involved in the particular transmission

channel used. For example, (Canini et al. 2014) have shown how category learning biases can emerge naturally through an iterated learning study. In this way, iterated learning experiments can be used as a technique to uncover the cognitive biases of participants.

However, to date, only a few studies have investigated the emergence of vocal iconicity through iterated learning. Jones et al. (2014) trained participants on miniature languages that consisted of pairings between various round and pointy shapes and written labels which were rated as sounding iconically neutral by English monolinguals. The participants then had to type the label learnt for each shape, including shapes they had not previously been trained on, and these labels were passed on to the next participant. Jones et al. found that iconic labels emerged to express round shapes but not pointy ones. When the participants then had to match labels that were judged as either iconically round, pointy or neutral to one of two shapes, they again only found an effect for the round shapes, which therefore suggested that the driving force behind this type of iconic mapping is the lip shape involved when producing round sounds rather than a cross-modal diagrammatic mapping.

Tamariz et al. (2017) conducted a similar study in which participants were assigned to one of two conditions. The first condition was a standard iterated learning design, as described above: participants had to learn the mapping between written non-words and meanings (spiky and round figures) and this mapping was then taught to a new participant, and so forth. In the second condition, there were two participants in each generation who used the words to communicate with each other. The authors found that the emergent words were rated as more pointy under the communicative condition, suggesting that the process of communicating with others contributes to stronger iconicity effects. Carr et al. (2017) also found that iconic patterning can emerge through iterated learning. In their experiments, participants had to learn words (presented in both written and auditory form) for randomly generated triangles. Although the study was not designed to investigate vocal iconicity directly, the authors nevertheless noted that thinner triangles tended to be labelled by sounds listed as 'pointy' by Ahlner and Zlatev (2010) (e.g. /k/, /i/, /t/), while more equilateral triangles tended to be labelled using sounds listed as "round" (e.g. /b/, /m/, /u/). They found this effect under both a standard iterated learning design and a design in which participants had to communicate. Furthermore, Edmiston et al. (2018) showed that when environmental sounds, such as breaking glass or splashing water, are imitated through iterated learning, they

become more stable and word-like, resembling ideo-phones. The final forms of the imitations could be matched to the source sounds above chance. Likewise, when people are asked to make up novel vocalizations for basic vocabulary words, naïve listeners are able to infer what they mean based on their phonetic forms (Parise and Pavani 2011; Perlman et al. 2015; Perlman and Lupyan 2018).

Based on previous studies, we know that the meaning pairs SMALL-BIG and POINTY-ROUND have been found to be consistently associated with sounds. These studies have also shown that a binary distinction between different types of sounds seems to be beneficial for studying oppositional vocal iconicity. We do not, however, know exactly which phonetic parameters are relevant for understanding iconicity in the SIZE- and SHAPE-domains, including broad categories such as vowels and consonants, since the premade stimuli words previously used have generally included a mix of sounds which belong to several different parameters. Hence, in order to understand how iconic associations emerge and which sounds and meanings are driving these effects, this study adopts a new approach for studying these phenomena using immediate iterated learning, which also bypasses the forced choice paradigm. Thus, this study adopts an explorative approach with the aim to reveal iconic correspondences. However, based on evidence from the large body of previous studies on sound-meanings associations in the SHAPE- and SIZE-domains, we could make some general assumptions. This included that 1, the meanings SMALL and POINTY, could result in words with a larger share of high or front vowels and consonants with high-frequency energy accumulation than the meanings BIG and ROUND, and 2, the meaning ROUND could result in words with a larger share of rounded vowels and labial consonants than the meaning POINTY.

2. Method

The methodological setup we used is relatively simple. The participants were divided into five conditions (CONTROL, BIG, SMALL, ROUND, and POINTY) and were presented with a recording of a single seed word, which includes a wide range of typologically common segments, and were asked to repeat it. These repetitions uttered by the participants were recorded and then used as stimuli for the next participant in the same transmission chain. This process was then repeated for 15 generations per transmission chain. In the CONTROL-condition, the word was simply passed down the 15 generations, but in the other conditions the participants were primed

with a meaning. The overall paradigm is illustrated in Fig. 1.

2.1 Participants

Participants were recruited online via the Figure Eight crowdsourcing platform which made it possible to include participants from several countries and with a range of different first languages. The participants were prevented from participating in the experiment more than once by identifying themselves with their unique worker IDs. The aim of the study was to include 15 generations (participants) per transmission chain and 20 transmission chains for each of the five conditions, for a total of 1,500 unique participants. To achieve this, we recruited 2,854 participants, 1,354 were of whom had to be excluded for one or more of the following reasons: 1, Misunderstanding the task, such as repeating the meaning ('big', 'small') rather than the word or asking a question about the task; 2, providing recordings of low quality (e.g. lack of sound, interfering background noise or recordings in which there were no recognizable sounds from the previous generation); or 3, providing recordings with obvious lexical interference, such as mistaking the presented audio as a word or phrase in a real language. The CONTROL-condition required 554 participants to yield 300 usable recordings, the BIG-condition required 592, the SMALL-condition required 591, the ROUND-condition required 565, and the POINTY-condition required 552. The participants were paid 50 cent USD for completing the task, which took around two minutes, and the study was conducted under established ethical standards approved by the Linguistics and English Language Ethics committee at the University of Edinburgh.

2.2 Stimuli

Of the five conditions, four were designed to prime the participants with a meaning by including either of the semantically oppositional poles of the SIZE-domain (BIG-SMALL) or the SHAPE-domain (ROUND-POINTY). The meanings for the BIG- and SMALL-conditions were conveyed in text since stimuli based on illustrations would require comparison in order to convey the correct meaning. The participants were either presented with the sentence 'The word you are going to hear means **big**' in the BIG-condition or with 'The word you are going to hear means **small**' in the SMALL-condition. The biases for the ROUND- and POINTY-condition were conveyed through shapes presented visually together with the sentence 'The word you are going to hear means', as shown in
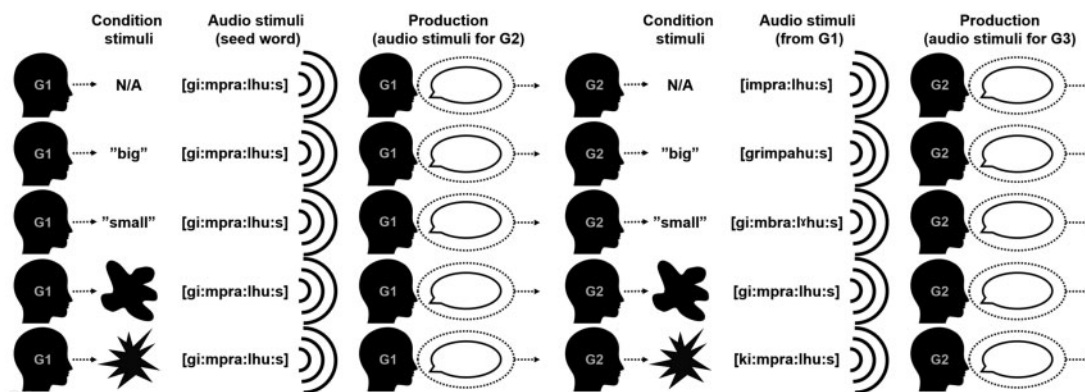
Figure 1. Illustration of the experimental procedure for the five conditions. The first-generation participants (G1) are exposed to their condition-specific visual stimuli and then to the seed word. They then repeat the word and their production was, in turn, used as the audio stimulus for the subsequent generation in the same transmission chain. This process was iterated until all chains had successfully transmitted the evolving string of sounds through 15 participants.
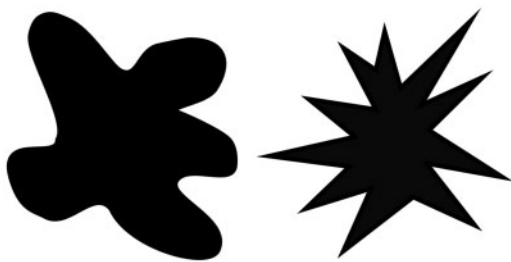

Figure 2. Visual stimuli for the ROUND- and POINTY-conditions.

Fig. 2. In the CONTROL-condition participants were not primed with a meaning.

All transmission chains were initialized with the same single seed word (i.e. the same audio stimulus was presented to the first participant in every transmission chain). This was to make it as easy as possible to track the development of sounds and groups of sounds over generations and for easier comparison across conditions. To allow for a variety of different potential iconic strategies to emerge, we designed the seed word to include a typologically, acoustically, and articulatory varied selection of segments.

It is difficult to ensure that each speech sound that a word contains will not result in any kind of semantic association for all speakers. This is not because all segments are iconic, but rather because of lexical transfer as a result of segments' varying occurrence in words across languages. In addition, there are also associations which could stem from the idiosyncratic salience that certain segments might have in an individual speaker's mental lexicon. We have neither a comprehensive

overview of all iconic mappings between sounds and meanings utilized throughout human languages, nor a list of potential language-specific or individual sound-meaning patterns. Thus, we need a seed word that is located at the center of cross-linguistic phonological space to allow for sound changes in any direction across generations of participants, especially in regard to speakers' different perception of speech sounds due to language-specific phonological systems. Furthermore, this word has to accommodate a reasonable mutation rate (i.e. to ensure that the seed word can evolve phonetically, it should be somewhat difficult to remember). If the word were too easily learned, the participants would be able to repeat it perfectly and there would hence be no space for evolution to operate in.

The seed word was designed to consist of three syllables. The sounds were selected to be present in more than half of the phonologies of the world's languages ([r] is the most common vibrant but was only found in 44% of the phonologies) (Mielke 2017; Moran and McCloy 2019) since the initial seed word was assumed to adapt to the participants' phonologies quickly which would leave the use of uncommon sounds for increasing mutation rates unnecessary. Long versions of the three most extreme vowels, [i :], [a :], and [u :], were included, and the seven featured consonants were selected to be evenly distributed across manners and positions of articulation, as shown in Table 1.

Approximately the same number of voiceless and voiced consonants was used in the word and consonant clusters were designed to include both voiced and voiceless sounds. In addition, the voiceless consonants were placed in the same syllables as the vowels with lower F_2,

Table 1. Distributions of consonants across five generalized manners of articulation and three generalized positions of articulation in the seed word (generation 0).

	Labial		Alveolar/palatal		Velar/glottal	
	−voice	+voice	−voice	+voice	−voice	+voice
Nasal		[m]				
Stop	[p]					[g]
Fricative			[s]		[h]	
Vibrant				[r]		
Lateral				[l]		

[u] and [a], and the voiced grave (Jakobson et al. 1953) consonants in the same syllable as the vowel with the lowest F_2, [i], to distribute the general spectral energy throughout the entire word. The selected parameters resulted in the word form [gi: mpra: lhu: s] which was then recorded by a female native speaker of Czech with an academic background in linguistics to ensure a phonetically accurate pronunciation of the word. The selected segments of the seed word are present in, on average, 76% of the 2155 phonologies available in the cross-linguistic phonological inventory repository PHOIBLE (Moran and McCloy 2019): [g] 64%, [i] 93%, [m] 95%, [p] 87%, [r] 38%, [a] 91%, [l] 66%, [h] 65%, [u] 87%, and [s] 77%.

2.3 Procedure

The task began with the following general instructions: 'In this task you will hear a word in an "alien" language. We will also tell you the meaning of the word. Your task is to listen carefully to the word and repeat it into your microphone. Make sure your speakers or headphones are switched on and the volume is turned up. First, we will tell you **the meaning of the word**. Then you will **hear the word**. There will then be a **3-second pause**. Finally, you must **repeat the word** into the microphone.' Participants in the CONTROL-condition, however, were not told that they would be presented with the meaning of the word.

Next, the participants entering the ROUND- and POINTY-conditions were presented with the round and pointy shapes. Those entering the BIG- and SMALL-conditions were presented with the text stimuli and were then required to confirm that they read the text properly by typing 'big' or 'small' depending on condition in order to continue with the task. This was included to make sure that the participants actually actively read the text stimuli since these could be easily overlooked as compared with the shape stimuli. This step was skipped for the participants in the CONTROL-condition who instead proceeded directly to the listening and production steps.

The first participant in each transmission chain listened once to the constructed seed word, which was followed by a 3-s pause after which they had to repeat what they heard into their microphone. After completing the task, the participants were asked what they thought the word meant along with a few background questions (native and other languages). The utterance that the participant recorded was then uploaded to our server. All recorded stimuli were manually checked by the experimenter. Often it was also necessary to normalize the volume to a consistent level and/or trim the recording to only include the actual utterance. The recorded utterance was then used as the stimulus for the next participant in the same transmission chain.

2.4 Data analysis

After data collection was completed (audio files can be accessed at https://osf.io/y3eru/, along with other supplementary material (see online), Appendix 4 and Appendix 5), the audio recordings were transcribed into the International Phonetic Alphabet (Appendix 1). This was done by the first author who was blind to the conditions the data belonged to. Tones, stress or phonemic length were not taken into consideration for the analysis as they seldom are transmitted correctly when speakers from different languages attempted to pronounce utterances with these features. Diphthongs, triphthongs, affricates and coarticulations were divided into their components and analyzed as separate segments for comparability reasons.

The transcribed sounds were then categorized according to six binary *sound parameters*, three for vowels and three for consonants (Appendix 5). Vowels were divided into HIGH and LOW, FRONT and BACK, and ROUNDED and UNROUNDED, while consonants were divided into GRAVE and ACUTE, VOICED and VOICELESS, and SONORANT and OBSTRUENT (see Table 2). The HIGH-group included high, near-high, high-mid and true-mid vowels (including [ə]), while the remaining vowels were assigned to the LOW-group. The FRONT-group included front and near-front vowels and the BACK-group included central, including [ə], near-back and back vowels. The ROUNDED-group included all rounded vowels and UNROUNDED-group unrounded vowels. Likewise, the VOICELESS-group included all voiceless consonants, VOICED-group all voiced consonants, the SONORANT-group all sonorant consonants and the OBSTRUENT-group all obstruent consonants. Finally, the GRAVE-group included bilabials through linguolabials, as well as velars through

Table 2. The included sound parameters and sound groups, as well as examples of typologically common segments of each sound group.

Principal class	Sound parameter	Sound group	Segment examples
Consonant	Voicing	VOICELESS	p, t, k, f, s, h
		VOICED	m, n, ŋ, b, d, g, v, z, w, l, r, j
	Manner	SONORANT	m, n, ŋ, w, l, r, j
		OBSTRUENT	p, t, k, b, d, g, f, s, h, v, z
	Position	GRAVE	m, ŋ, p, k, b, g, f, h, v, w
		ACUTE	n, t, d, s, z, l, r, j
Vowel	Height	HIGH	i, e, ə, u
		LOW	a, o
	Backness	FRONT	i, e, a
		BACK	ə, o, u
	Roundedness	UNROUNDED	i, e, a, ə
		ROUNDED	o, u

glottals, and the ACUTE-group included dentals through palatals.

2.5 Statistical model

We modeled the proportion of vowels or consonants of each particular sound parameter (HIGH-LOW, FRONT-BACK, ROUNDED-UNROUNDED, GRAVE-ACUTE, VOICED-VOICELESS, SONORANT-OBSTRUENT) out of the total number of vowels or consonants in the word for generation 0 (seed word) through 15. Proportions rather than absolute values were chosen in order to compensate for reduplication and word length effects. The proportions were calculated separately for vowels and consonants since it is possible that some transmission chains might utilize the former iconically, while others might utilize the latter. For example, if an association is found between a meaning and high frequency sounds, the sound could be voiceless consonants, front unrounded vowels, or both. Thus, a phonetic form such as [tuta] was analyzed as 100% [t] in terms of its consonants, and 50% [a] and 50% [u] in terms of its vowels. We then used binomial mixed models with generation and condition as predictors, with an interaction. One such model was fit for each of the six sound parameters. To account for the non-independent nature of observations from the same chain, we included chain as a random intercept. This may mitigate the problem of autocorrelation of residuals from adjacent observations. To minimize the risk of overfitting, we imposed their conservative shrinkage to zero with the horseshoe prior (Carvalho et al. 2009). The models were fit using R 3.6.0 (R Core Team 2020) and the package *brms* version 2.9.0 (Bürkner 2017). We first modeled the changes in proportion of each sound parameter and condition, including the CONTROL-

condition. We then also compared the changes of proportions for each of the stimuli-conditions to the changes of proportions of the CONTROL-condition.

3. General results

In total, the participants reported 58 different first languages which can be found in Appendix 4. Two thirds of all participants reported one of the five most common languages: Spanish (485), English (223), Serbo-Croatian (104), Russian (99), and Arabic (88). On average, the original 10 segments (3 vowels, 7 consonants) of the seed word were reduced by approximately 3 at generation 15, as seen in Fig. 3.

The reduction of total word length was mainly caused by the loss of consonants, which at generation 15 were reduced from the original 7 to approximately 4. The vowels, on the other hand, were only reduced by about a quarter of a segment on average. It is quite possible that the reason for these differences between consonants and vowels could be attributed to a general articulatory preference for simple syllable structures, see for example the example transmission chains in Table 3.

First, we tested whether we could find any noteworthy over- or underrepresentations of the sound groups when comparing the seed word to the generation 15 words within each condition. Since all sounds groups were constructed in pairs, an overrepresentation of, for example, rounded vowels would correspondingly also result in an underrepresentation of unrounded vowels. All conditions, except BIG, showed noteworthy changes for at least two of the investigated sound parameters (see Fig. 4 and Appendix 2). However, the HIGH-LOW and SONORANT-OBSTRUENT parameters did not produce any noteworthy changes. The proportion of FRONT vowels
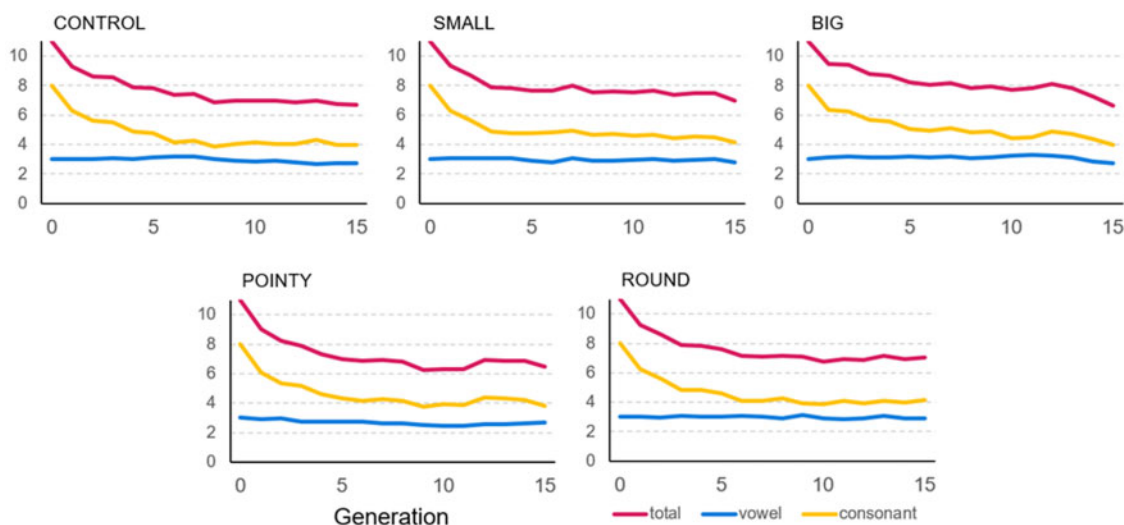
Figure 3. Change in total, vowel and consonant segment length over generations.

Table 3. An example transmission chain for each of the conditions which shows the phonetic changes from generation 0 (seed word) to generation 15.

Generation	CONTROL	SMALL	BIG	POINTY	ROUND
0	grimpralhus	grimpralhus	grimpralhus	grimpralhus	grimpralhus
1	gimpralhus	gɹimbɹahahu	grimpralhus	imprashus	gimpɹalɹus
2	gimrahut	nɛŋhaahu	grinkraprus	imprasus	infraru
3	gimrahu	mɛjaahu	gringrasgrʊs	intraʊs	infraru
4	igaɹʊ	nihahahu	gringranrʊs	intrabus	infagʊw
5	igarʊ	hahu	wewiwagʊs	içabow	difaigum
6	igaru	jahu	wewiwaŋwʊs	izabo	ifogo
7	egaru	jahu	weʈiʈaŋgos	izabo	ifogo
8	egarʊ	jahu	weʈhihaho	bizabo	ifogo
9	egarʊ	jahu	weʁhihaʁho	isawo	ifogo
10	ejaɹʊ	jahʊ	ewibihaʁbiχ	isawo	ifodu
11	ejalʊk	bjahu	ivivibixabixa	isawok	infodo
12	ejalʊk	mjahu	avirbihabiha	isawok	infodo
13	ejalʊk	wiaɹhu	ɛrbirlihæ	pisaloko	inforðo
14	ejlu	hiaɹhow	ɛnəmbjurixat	susaloko	okordo
15	ejlo	siathaw	inembuɹiha	tusalopo	hokogo

decreased in the CONTROL-condition (−13.2% 95% CI [−20.7, −5.1]), the POINTY-condition (−7.1% [−16.1, −0.1]) and the ROUND-condition (−13.7% [−21.6, −6]). Correspondingly, the proportion of ROUNDED vowels increased in the ROUND-condition 12.2% [3.6, 19.8] which is to be expected, since, typologically, rounded vowels are generally back while unrounded vowels are front. Conversely, the SMALL-condition produced a notable decrease of ROUNDED vowels (−10.3% [−16.9, −2.8]). The proportion of GRAVE consonants decreased in all

conditions; CONTROL-condition (−7.1% [−13.9, −0.7]), SMALL-condition (−14% [−19.3, −8.3]), BIG-condition (−15.1% [−21.6, −9.3]), POINTY-condition (−15.8% [−21.4, −10.6]), ROUND-condition (−8.1% [−14.4, −0.8]). Lastly, the proportion of VOICED consonants increased slightly in ROUND-condition (7.3% [1.8, 13]).

Secondly, we compared the sound distributions for each stimuli condition to the CONTROL-condition (see Fig. 5 and Appendix 3). This crystalized the results and became easier to interpret. There were two cases for which
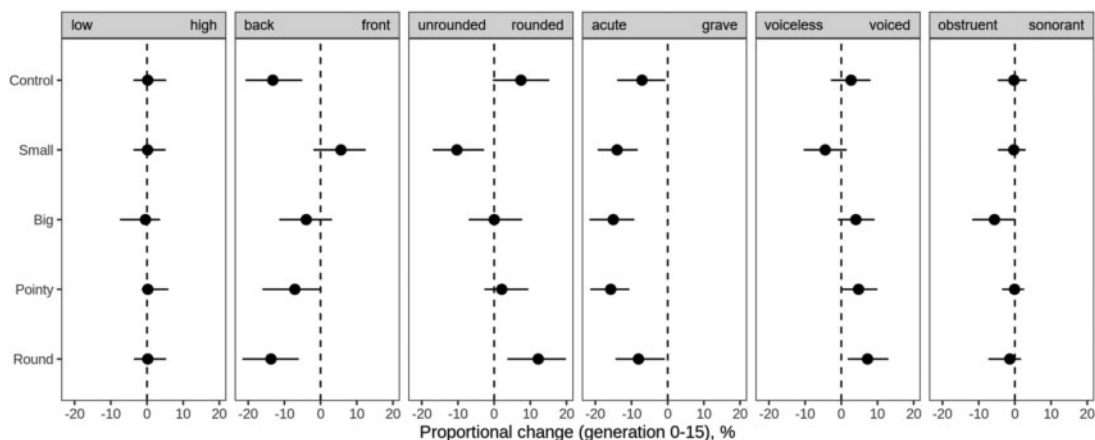
Figure 4. Change in the proportion for the six sound parameters from generation 0 to generation 15. Shown: median of posterior distribution and 95% CI. Values which are completely on either side of the vertical dashed line were considered statistically reliable.

the 95% CI clearly excluded zero. First, the proportion of FRONT vowels increased in the SMALL-condition by an additional 18.8% [8.3, 27.9] compared to the CONTROL-condition. Second, this was mirrored by a decrease of the proportion of ROUNDED vowels in the SMALL-condition versus the CONTROL-condition by −17.8% [−27.0, −7.4]. In addition, a weaker yet noteworthy effect was found for the proportion of GRAVE consonants which decreased in the POINTY-condition by −8.7% [−16.6, −0.5] when compared to the CONTROL-condition.

Furthermore, as shown in Fig. 6, the noteworthy changes compared to the CONTROL-condition started taking off around generation 6 and gradually increased, which can be seen most clearly in the rounded-unrounded parameter. This could be attributed to word length and syllabic complexity which might have distracted the participants from the text and shape stimuli. The average decrease of word length (four sounds) was most prominent in the early generations; by generation 6 the word lengths had decreased by three sounds and by generation 15 the word lengths had only decreased by one additional sound. Thus, the stimuli words would have had to become simplified before iconicity could start affecting the sounds in the words. However, this suggests that it is possible that even stronger effects might be observed over longer transmission chains (cf. Tamariz et al. 2017).

4. Discussion

The purpose of this study was to investigate how iconic associations emerge and which sounds and meanings are involved by observing how a single seed word was altered by being transmitted between language users in five different conditions.

4.1. General discussion

The most important results were yielded by comparisons between the CONTROL-condition and the stimuli conditions, iconic effects were found for both vowels and consonants. This is also in line with other studies that have shown that both vowels and consonants are involved in size and shape iconicity (Ahlner and Zlatev 2010; Nielsen and Rendall 2013; D'Onofrio 2014). The clearest results were produced by the SMALL-condition and showed a preference for FRONT and UNROUNDED vowels and a dispreference for BACK and ROUNDED vowels. The preferred sounds were typically represented by [i], [e], [ɛ], and [a] which also have the highest average vowel frequencies for the first formant ([a] and [ɛ]) and for the second formant ([i] and [e]). Thus, the associations between sound and meaning align well with Ohala's (1994) frequency code which predicts that smallness, as well as related meanings, are evoked by high and/or rising frequencies of vocalizations. Furthermore, a plethora of cross-linguistic and experimental studies have found similar associations between size and energy level or pitch (Sapir 1929; Newman 1933; Taylor and Taylor 1962; Fónagy 1963; Taylor 1963). For example, Erben Johansson et al. (2020) found SMALL and SHORT to be associated with voiceless consonants, which of course also involve high frequency energy (Ohala 1994). Consequently, this association should probably be regarded as one of the most robust iconicity effects
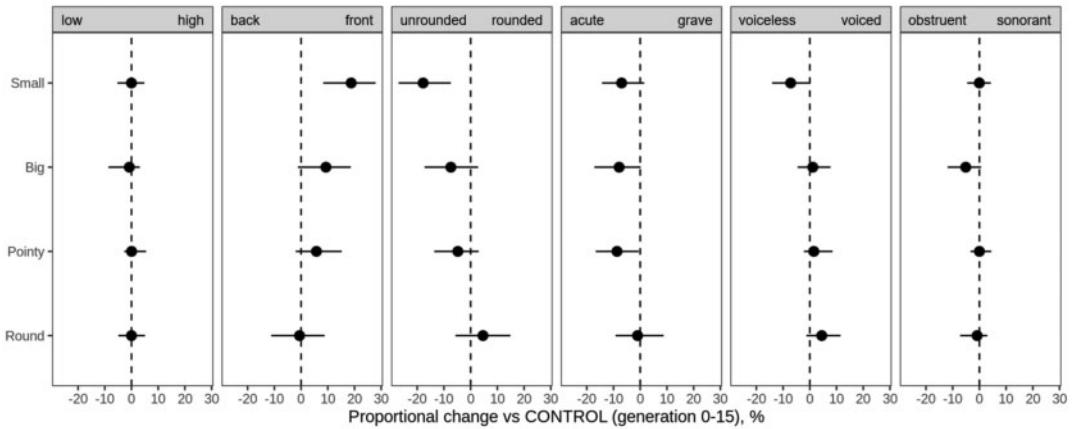
Figure 5. Contrasts between each stimuli-condition and the CONTROL-condition in the change in proportion for the six sound parameters from generation 0 to generation 15. Shown: median of posterior distribution and 95% CI. Values which are completely on either side of the vertical dashed line were considered statistically reliable.
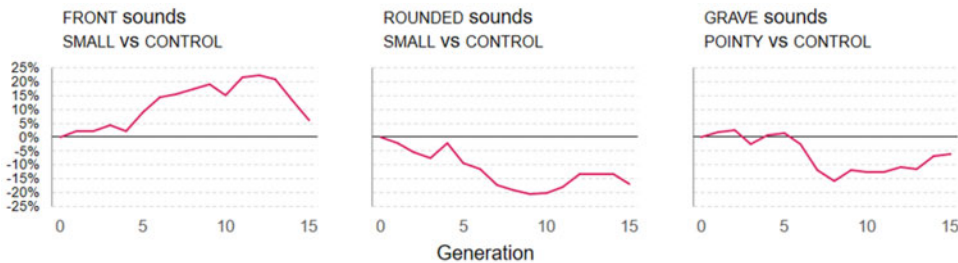
Figure 6. Showing the average proportional change of the three sound parameters which were found to be noteworthy when compared to the CONTROL-condition from generation 0 to generation 15. Left: Proportional change of the FRONT sound group (and reversely the BACK sound group) in the SMALL-condition vs the CONTROL-condition. Center: Proportional change of the ROUNDED sound group (and reversely the UNROUNDED sound group) in the SMALL-condition vs the CONTROL-condition. Right: Proportional change of the GRAVE sound group (and reversely the ACUTE sound group) in the SMALL-condition vs the CONTROL-condition.

found since it aligns with solid typological and experimental evidence.

The most surprising result was the decrease of GRAVE consonants, and the corresponding increase of ACUTE consonants in the POINTY-condition, since one of the most common GRAVE consonants, [k], is often featured in pointy stimuli words (e.g. [kiki]). The results do, however, align with the idea that consonants might play a somewhat larger role than vowels in shaping vocal iconicity (Nielsen and Rendall 2011; Fort et al. 2015). This does not necessarily mean that [k] is confirmed to be disfavored when paired with pointy shapes, since the sound group also contains labial and voiced consonants. Nevertheless, this has some implications for bouba-kiki tasks since it demonstrates that using ready-made stimuli words for experiments such as this might not always produce accurate effects (Dingemanse et al. 2015).

These findings also suggest a slightly more complex mapping between sound and meaning than pitch-to-size. Although no effect was found for the VOICED-VOICELESS parameter, ACUTE sounds do generally involve higher frequency energy than grave sounds, but this sound group included both voiceless and voiced sounds which is the primary consonantal distinction between high and low frequency energy. It is thus possible that the sharpness produced by ACUTE sounds are more fundamental than the overall energy differences between VOICED and VOICELESS sounds (see also Aryani et al. 2020). Consequently, there might also be a potential discrepancy between associated phonetic parameters and semantic domains. As one pole of the continuous SIZE-domain, the SMALL-condition can be clearly linked to the equally continuous frequency scale, but the sounds mapped to POINTY-condition could be, at least partially,

based on some other, potentially dichotomous, type of mapping. The preference for these different types of associations could be grounded in the semantic features of the stimuli as BIG and SMALL are rather abstract and require comparison in order to be defined which is a good fit for degrees of pitch. ROUND and POINTY are considerably more visually concrete and their contrasting geometrical features could also be used to tell them apart from shapes such as squares or ellipses. Accordingly, the sounds associated with POINTY could portray similar concreteness. Alternatively, the differences in presentation between the text stimuli in the SMALL- and BIG-conditions and shape stimuli in the POINTY- and ROUND-conditions might have resulted in the conditions not being completely comparable due to different confounding factors. Nevertheless, it is unlikely that this would completely explain why vowels were found to be associated with SMALL and consonants with POINTY. It is also possible that there could be a specific orthographic confound in the SMALL- and BIG-conditions (Cuskley et al. 2017). While no effect was found for the BIG-condition when compared to the CONTROL-condition, the SMALL-condition did include an/a/which represents an unrounded front vowel in many languages and could have led to an increase of similar sounds in the results. However, in the English pronunciation of the word *small*, /a/ represents /ɔ/, a rounded back vowel which could indicate that this effect would be rather modest.

Another interesting finding was that, when compared to the CONTROL-condition, the SMALL- and POINTY-conditions produced iconic effects while the BIG- and ROUND-conditions did not. It is difficult to know exactly why only one of the poles of these semantic parameters, but similar results have been found in previous studies although both for pointy and for round shapes (Nielsen and Rendall 2011; Jones et al. 2014; Tamariz et al. 2017; Fort et al. 2018). There is also evidence for that antonyms and semantically oppositional concepts are cognitively closely related (Deese 1965; Justeson and Katz 1991; Willners 2001; Paradis et al. 2009), that poles of the same semantic dimension differ in their iconic predictability (Westbury et al. 2018) and that iconic relationships between concepts can be upheld in reversed order (Johansson and Carling 2015). This could therefore indicate that these types of concepts are iconically coded pairwise.

Finally, it should also be pointed out that the CONTROL-condition produced two noteworthy changes and all stimuli conditions produced decreases in GRAVE consonants which illustrates the difficulty with designing the seed word. However, the decrease in front vowels and increase in rounded vowels in the CONTROL-condition are in fact mirror images of each other since front vowels are usually unrounded. Furthermore, both of these effects are found in the experimental conditions as well, with the notable exception of the SMALL-condition. In addition, to minimize the risk of finding effects by chance, we controlled for multiple comparisons by imposing a conservative shrinkage prior (see Section 2.5). This suggests that these changes could be interpreted as a stabilization toward a kind of typological and/or articulatory default. Furthermore, it could be assumed that the proportions of iconic sounds would increase indefinitely until the transmitted words would consist only of front unrounded vowels. This is, however, unlikely for a number of reasons since linguistic material from various sources is dynamically introduced into words as languages change over time. First, words, except for a very small number per language, generally adhere to phonotactic restrictions that require them to include both vowels and consonants. This is because there simply are not enough unique individual phonemes in languages to be assigned to all meanings that need to be conveyed. And secondly, many languages require all words, including loans, to have affixes attached to them in order to be grammatical. Similarly, the participants included in the present study were also instructed to repeat what they heard which forced them to retain considerable parts of the syllable structure and sounds from the previous utterance.

4.3 What is required for iconicity to emerge?

Jones et al. (2014) showed that iconicity can emerge through transmission. However, as with most previous experiments that have investigated iconicity, the participants were highly restricted due to the use of text-based artificial languages or forced-choice experimental design. While accompanied by the same methodological restrictions, Tamariz et al. (2017) only found that iconicity emerges through communicative interaction and not through individual reproduction. The stronger effect that interaction brings to the table was attributed to an increased number of possible innovations that could increase iconicity as well as a larger number of possible adopters of the signal, which increases the chance of labels fitting with meanings in a speech community. Therefore, Tamariz et al. argue that their results can be interpreted as evidence for random mutation and selection rather than guided variation; in other words, cultural traits acquired by a population through individual learning drive cultural evolutionary processes.

Furthermore, the overall reason or reasons for how iconicity emerges over time is harder to tease apart and

outside of this article's scope. However, it can be assumed that, in the present study, either the participants' memories were affected causing them to misremember a more iconic version of the stimulus, or their perceptions were affected causing them to perceive a more iconic version of the stimulus. Alternatively, this could also entail a combination of both of these, or more, factors, such as recordings being muffled which created vaguer stimuli that would give rise to more change and assimilation of the phonemes of the stimuli words into the native phonology of the participants. This could therefore somewhat limit the generalizability of the present study.

Nonetheless, it would be unwise to underestimate the role of transmission in the dynamics of iconicity. First, both Jones et al. (2014) and the present study showed that transmission alone is enough for iconic effects to arise. Secondly, the present study further suggests that very little is required in order for iconicity to emerge (Edmiston et al. 2018). Even without interaction between participants, constrained experimental setups, forced choice questions, premade stimuli words or using text as a proxy for spoken language, all of which could in some manner increase the likelihood of mapping sound to meanings correctly outside of the bouba-kiki effect (Cuskley et al. 2017), iconic effects seem to have emerged. Thirdly, there is overwhelming evidence that iconic forms, including language-specific ideophones, facilitate language learning and comprehension in both children and adults (Imai et al. 2008; Nygaard et al. 2009; Kantartzis et al. 2011; Imai and Kita 2014; Lockwood et al. 2016a,b; Massaro and Perlman 2017).

However, while iconicity and synesthetic cross-modal mappings are present in the early stages of human ontogenetic development (Mondloch and Maurer 2004; Maurer et al. 2006; Walker et al. 2010) and go at least as far back as to the ancestor we share with chimpanzees (Ludwig et al. 2011; Perlman 2017), a recent study failed to find a bouba/kiki effect in great apes (Margiotoudi et al. 2019). In addition, these mappings do not seem to disappear but as language competence and vocabulary size increases with age, the share of iconicity in the lexicon is lower for adults as compared to children (Ludwig and Simner 2013; Perry et al. 2015; Massaro and Perlman 2017). The likely explanation for this is that iconicity does not scale well in language. In a less developed and lexically poor language, iconicity can aid in intuitively linking words to fundamental meanings, but as languages adapt to the expressive needs of their users, the number of distinctions that must be made cannot be handled by an iconic system. Thus, here is where iconicity falls short, as there simply are not enough unique iconic signals (either through sounds or gestures) available to accommodate the diversity of meanings that language users might wish to express (Gasser 2004; Westbury et al. 2018).

Nevertheless, iconicity is still found in complex languages and seems to permeate many sections of the lexicon (Sidhu et al. 2019), although it excels in specific functions in conjunction with arbitrary and systematic mappings between sound and meaning (Monaghan et al. 2011; Dingemanse et al. 2015). However, agents without advanced language competence, such as great apes, do utilize iconicity even though they have very limited access to interactional language, which suggests that the transmission of signals is enough to facilitate iconicity. Furthermore, large-scale cross-linguistic studies on lexical iconicity show that iconic forms are present throughout languages and language families, but also that the same sound-meaning associations are not found everywhere at the same time (Wichmann et al. 2010; Blasi et al. 2016; Erben Johansson et al. 2020; Joo 2020), which suggests that iconicity is in a perpetual process of decay and rebuilding (Johansson and Carling 2015) and not conserved through time (Pagel et al. 2013). In sum, certain iconic associations between speech sounds and semantic features seem to affect the formation of lexemes in human language and while interaction could provide an even more advantageous environment for non-arbitrary associations, this study suggests that interaction might not be a prerequisite for iconicity.

5. Conclusion

We have shown that by adopting an iterated learning approach for investigating the classic bouba-kiki and mil-mal experiments, as well as including a much larger number of participants than previous studies, it was possible to see iconic effects emerge. By using a simple methodological setup which included an auditorily modest linguistic environment without premade stimuli words or task training, we were able to get a deeper understanding of how vocal iconicity operates within the semantic SIZE- and SHAPE-domains. Not only were these results aligned with the sound-meaning associations found in large-scale cross-linguistic and experimental studies, but one of the effects gradually strengthened with generation as well, which indicates that stronger effects might be observed with longer transmission chains. Furthermore, only the SMALL- and POINTY-conditions produced iconic effects, while the BIG- and ROUND-conditions did not, which could indicate that iconic effects do not have to be equal in concepts belonging to the same semantically oppositional pair. In addition, while the sounds associated

with the SMALL-condition could be linked to differing degrees of frequency, the sounds associated with the POINTY-condition could indicate that another factor serves as a foundation for this mapping and should be investigated further. In sum, these results indicate that linguistic transmission through disconnected language users is enough to investigate cognitive biases for vocal iconicity, which can easily be expanded to a range of iconically promising meanings, for example TALL, LONG, MANY, etc., and is of importance for our understanding of how iconicity emerges and decays, and how it can shape lexicons.

Supplementary data

Acknowledgements

References

Ahlner, F., and Zlatev, J. (2010). *Cross-Modal Iconicity: A Cognitive Semiotic Approach to Sound Symbolism*. https://doi.org/10.12697/SSS.2010.38.1-4.11

Anikin, A., and Johansson, N. (2019) 'Implicit Associations between Individual Properties of Color and Sound', *Attention, Perception, & Psychophysics*, 81/3: 764–77.

Aryani, A., Isbilen, E. S., and Christiansen, M. H. (2020) 'Affective Arousal Links Sound to Meaning', *Psychological Science*, 31/8: 978–86.

Blasi, D. E. et al. (2016) 'Sound–Meaning Association Biases Evidenced across Thousands of Languages', *Proceedings of the National Academy of Sciences*, 113/39: 10818–23.

Bross, F. (2018) 'Cognitive Associations between Vowel Length and Object Size: A New Feature Contributing to a Bouba-Kiki Effect', *Proceedings of the Conference on Phonetic & and Phonology in German-Speaking Countries*, 13: 17–20.

Bürkner, P.-C. (2017) 'Brms: An R Package for Bayesian Multilevel Models Using Stan', *Journal of Statistical Software*, 80/1: 1–28.

Canini, K. R. et al. (2014) 'Revealing Human Inductive Biases for Category Learning by Simulating Cultural Transmission', *Psychonomic Bulletin & Review*, 21/3: 785–93.

Carling, G., and Johansson, N. (2014) 'Motivated Language Change: Processes Involved in the Growth and Conventionalization of Onomatopoeia and Sound Symbolism', *Acta Linguistica Hafniensia*, 46/2: 199–217.

Carr, J. W. et al. (2017) 'The Cultural Evolution of Structured Languages in an Open-Ended, Continuous World', *Cognitive Science*, 41/4: 892–923.

Carvalho, C. M., Polson, N. G., and Scott, J. G. (2009) 'Handling Sparsity via the Horseshoe', *Proceedings of the 12th International Conference on Artificial Intelligence and Statistics*, 5: 73–80

Cuskley, C. (2013) 'Mappings between Linguistic Sound and Motion', *Public Journal of Semiotics*, 5/1: 39–62.

——, and Kirby, S. (2013). 'Synesthesia, Cross-Modality, and Language Evolution'. In: Simner, J. and Hubbard E. M. (eds.) *The Oxford Handbook of Synesthesia*, pp. 869–99. Oxford: Oxford University Press.

——, Simner, J., and Kirby, S. (2017) 'Phonological and Orthographic Influences in the Bouba-Kiki Effect', *Psychological Research*, 81/1: 119–30.

Deese, J. (1965). *The Structure of Associations in Language and Thought*. Baltimore: The Johns Hopkins Press, pp. Ix + 216.

Dingemanse, M. (2011). 'Ezra Pound among the Mawu'. In: Michelucci P., Fischer O., and Ljungberg C. (eds.) *Semblance and Signification. Iconicity in Language and Literature*, Vol. 10. Amsterdam: John Benjamins Publishing Company, pp. 39–54.

—— et al. (2015) 'Arbitrariness, Iconicity, and Systematicity in Language', *Trends in Cognitive Sciences*, 19: 603–15.

D'Onofrio, A. (2014) 'Phonetic Detail and Dimensionality in Sound-Shape Correspondences: Refining the Bouba-Kiki Paradigm', *Language and Speech*, 57/3: 367–93.

Edmiston, P., Perlman, M., and Lupyan, G. (2018) 'Repeated Imitation Makes Human Vocalizations More Word-Like', *Proceedings of the Royal Society B: Biological Sciences*, 285/1874: 20172709.

Erben Johansson, N. et al. (2020) 'The Typology of Sound Symbolism: Defining Macro-Concepts via Their Semantic and Phonetic Features', *Linguistic Typology*, 24/2: 253–310.

Fónagy, I. (1963). *Die Metaphern in Der Phonetik: Ein Beitrag Zur Entwicklungsgeschichte Des Wissenschaftlichen Denkens*. The Hague: Mouton & Co.

Fort, M. et al. (2018) 'Symbouki: A Meta-Analysis on the Emergence of Sound Symbolism in Early Language Acquisition', *Developmental Science*, 21/5: e12659.

——, Martin, A., and Peperkamp, S. (2015) 'Consonants Are More Important than Vowels in the Bouba-Kiki Effect', *Language and Speech*, 58/2: 247–66.

Gasser. (2004) 'The Origins of Arbitrariness in Language', *Proceedings of the Annual Meeting of the Cognitive Science Society*, 26: 434–9

Gebels, G. (1969) 'An Investigation of Phonetic Symbolism in Different Cultures', *Journal of Verbal Learning & Verbal Behavior*, 8/2: 310–2.

Goddard, C., and Wierzbicka, A. (2002). *Meaning and Universal Grammar*. Amsterdam: John Benjamins Publishing Company.

Hamilton-Fletcher, G. et al. (2018) 'The Role of Visual Experience in the Emergence of Cross-Modal Correspondences', *Cognition*, 175: 114–21.

Hinton, L., Nichols, J., and Ohala, J. J. (1994). 'Introduction: Sound-Symbolic Processes'. In: Hinton L., Nichols J., and Ohala J. J. (eds.) *Sound Symbolism*. Cambridge: Cambridge University Press, pp. 325–47.

Holland, M. K., and Wertheimer, M. (2016). Some Physiognomic Aspects of Naming, or, Maluma and Takete Revisited. *Perceptual and Motor Skills*. https://doi.org/10.2466/pms.1964.19.1.111.

Huang, Y.-H., Pratoomraj, S., and Johnson, R. C. (1969) 'Universal Magnitude Symbolism', *Journal of Verbal Learning & Verbal Behavior*, 8/1: 155–6.

Imai, M., and Kita, S. (2014) 'The Sound Symbolism Bootstrapping Hypothesis for Language Acquisition and Language Evolution', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369/1651: 20130298.

——— et al. (2008) 'Sound Symbolism Facilitates Early Verb Learning', *Cognition*, 109/1: 54–65.

Jakobson, R., Fant, G., and Halle, M. (1953). *Preliminaries to Speech Analysis: The Distinctive Features and Their Correlates*. Cambridge, MA: The MIT Press.

Johansson, N., Anikin, A., and Aseyev, N. (2019) 'Color Sound Symbolism in Natural Languages', *Language and Cognition*, 12/1: 56–83.

———, ———, and Carling, G. (2015) 'The de-Iconization and Rebuilding of Iconicity in Spatial Deixis', *Acta Linguistica Hafniensia*, 47/1: 4–32.

Johnson, R. C. (1967) 'Magnitude Symbolism of English Words', *Journal of Verbal Learning & Verbal Behavior*, 6/4: 508–11.

Jones, J. M. et al. (2014) 'The Bouba Effect: Sound-Shape Iconicity in Iterated and Implicit Learning'. In: Bello, P. Guarini, M., McShane, M., and Scassellati, B. (eds) *Proceedings of the 36th Annual Meeting of the Cognitive Science Society*. Austin, TX: Cognitive Science Society, pp. 1114–1119.

Joo, I. (2020) 'Phonosemantic Biases Found in Leipzig-Jakarta Lists of 66 Languages', *Linguistic Typology*, 24/1: 1–12.

Justeson, J., and Katz, S. (1991) 'Co-Occurrence of Antonymous Adjectives and Their Contexts', *Computational Linguistics*, 17: 1–19

Kantartzis, K., Imai, M., and Kita, S. (2011) 'Japanese Sound-Symbolism Facilitates Word Learning in English-Speaking Children', *Cognitive Science*, 35/3: 575–86 .

Kirby, S. (2001) 'Spontaneous Evolution of Linguistic Structure-an Iterated Learning Model of the Emergence of Regularity and Irregularity', *IEEE Transactions on Evolutionary Computation*, 5/2: 102–10.

———, Cornish, H., and Smith, K. (2008) 'Cumulative Cultural Evolution in the Laboratory: An Experimental Approach to the Origins of Structure in Human Language', *Proceedings of the National Academy of Sciences*, 105/31: 10681–6.

———, and Hurford, J. R. (2002). 'The Emergence of Linguistic Structure: An Overview of the Iterated Learning Model'. In: Cangelosi A. and Parisi D. (eds.), *Simulating the Evolution of Language*. New York: Springer, pp. 121–147.

——— et al. (2015) 'Compression and Communication in the Cultural Evolution of Linguistic Structure', *Cognition*, 141: 87–102.

Köhler, W. (1929). *Gestalt Psychology*. Liveright, pp. xi, 403. [AU: Please provide publisher location for Köhler 1929.]

Ladefoged, P. (2001). *Vowels and Consonants: An Introduction to the Sounds of Languages*. Hoboken, NJ: Blackwell Publishing.

———, and Maddieson, I. (1996). *The Sounds of the World's Languages*. Hoboken, NJ: Blackwell Publishers.

Lapolla, R. (1994). 'An Experimental Investigation into Sound Symbolism as It Relates to Mandarin Chinese'. In: Hinton L., Nichols J., and Ohala J. J. (eds.), *Sound Symbolism*. Cambridge: Cambridge University Press, pp. 325–47.

Lindauer, M. S. (1990) 'The Meanings of the Physiognomic Stimuli Taketa and Maluma', *Bulletin of the Psychonomic Society*, 28/1: 47–50.

Lockwood, G., and Dingemanse, M. (2015) 'Iconicity in the Lab: A Review of Behavioral, Developmental, and Neuroimaging Research into Sound-Symbolism', *Frontiers in Psychology*, 6:1624.

———, ———, and Hagoort, P. (2016a) 'Sound-Symbolism Boosts Novel Word Learning', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42/8: 1274–81.

———, ———, ———, and Dingemanse, M. (2016b) 'How Iconicity Helps People Learn New Words: Neural Correlates and Individual Differences in Sound-Symbolic Bootstrapping', *Collabra: Psychology*, 2/1: 7.

Ludwig, V. U., Adachi, I., and Matsuzawa, T. (2011) 'Visuoauditory Mappings between High Luminance and High Pitch Are Shared by Chimpanzees (Pan troglodytes) and Humans', *Proceedings of the National Academy of Sciences*, 108/51: 20661–5.

———, ———, and Simner, J. (2013) 'What Colour Does That Feel? Tactile—Visual Mapping and the Development of Cross-Modality', *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, 49/4: 1089–99.

Maglio, S. J. et al. (2014) 'Vowel Sounds in Words Affect Mental Construal and Shift Preferences for Targets', *Journal of Experimental Psychology: General*, 143/3: 1082–96.

Margiotoudi, K. et al. (2019) 'Sound Symbolic Congruency Detection in Humans but Not in Great Apes', *Scientific Reports*, 9/1: 12705.

Massaro, D. W., and Perlman, M. (2017) 'Quantifying Iconicity's Contribution during Language Acquisition: Implications for Vocabulary Learning', *Frontiers in Communication*, 2: 1–14.

Maurer, D., Pathman, T., and Mondloch, C. J. (2006) 'The Shape of Boubas: Sound-Shape Correspondences in Toddlers and Adults', *Developmental Science*, 9/3: 316–22.

Mielke, J. (2017). PBase: A database of phonological patterns. Accessed at https://pbase.phon.chass.ncsu.edu on 12 April 2017.

Monaghan, P., Christiansen, M. H., and Fitneva, S. A. (2011) 'The Arbitrariness of the Sign: Learning Advantages from the Structure of the Vocabulary', *Journal of Experimental Psychology: General*, 140/3: 325–47.

Mondloch, C. J., and Maurer, D. (2004) 'Do Small White Balls Squeak? Pitch-Object Correspondences in Young Children', *Cognitive, Affective, & Behavioral Neuroscience*, 4/2: 133–6.

Moran, S., and McCloy, D. (eds) (2019). *PHOIBLE 2.0. Jena: Max Planck Institute for the Science of Human History*. Accessed at http://phoible.org on 12 April 2017.

Newman, S. S. (1933) 'Further Experiments in Phonetic Symbolism', *The American Journal of Psychology*, 45: 53–75.

Nielsen, A., and Rendall, D. (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 Experimentale*, 65/2: 115–24.

———, and ——— (2012) 'The Source and Magnitude of Sound-Symbolic Biases in Processing Artificial Word Material and Their Implications for Language Learning and Transmission', *Language and Cognition*, 4/2: 115–25.

———, and ——— (2013) 'Parsing the Role of Consonants versus Vowels in the Classic Takete-Maluma Phenomenon', *Canadian Journal of Experimental Psychology = Revue Canadienne De Psychologie Experimentale*, 67/2: 153–63.

Nygaard, L. C., Cook, A. E., and Namy, L. L. (2009) 'Sound to Meaning Correspondences Facilitate Word Learning', *Cognition*, 112/1: 181–6.

O'Boyle, M. W., and Tarte, R. D. (1980) 'Implications for Phonetic Symbolism: The Relationship between Pure Tones and Geometric Figures', *Journal of Psycholinguistic Research*, 9/6: 535–44.

Ohala, J. J. (1994). 'The Frequency Code Underlies the Sound Symbolic Use of Voice Pitch'. In: Hinton L., Nichols J., and Ohala J. J. (eds.), *Sound Symbolism*. Cambridge: Cambridge University Press, pp. 325–47.

Pagel, M. et al. (2013) 'Ultraconserved Words Point to Deep Language Ancestry across Eurasia', *Proceedings of the National Academy of Sciences of the United States of America*, 110/21: 8471–6.

Paradis, C., Willners, C., and Jones, S. (2009) 'Good and Bad Opposites: Using Textual and Experimental Techniques to Measure Antonym Canonicity', *The Mental Lexicon*, 4/3: 380–429.

Parise, C. V., and Pavani, F. (2011) 'Evidence of Sound Symbolism in Simple Vocalizations', *Experimental Brain Research*, 214/3: 373–80.

Perlman, M. (2017) 'Debunking Two Myths against Vocal Origins of Language: Language is Iconic and Multimodal to the Core', *Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems*, 18/3: 376–401.

———, Dale, R., and Lupyan, G. (2015) 'Iconicity Can Ground the Creation of Vocal Symbols', *Royal Society Open Science*, 2/8: 150152.

———, and Lupyan, G. (2018) 'People Can Create Iconic Vocalizations to Communicate Various Meanings to Naïve Listeners', *Scientific Reports*, 8/1: 2634.

Perry, L. K., Perlman, M., and Lupyan, G. (2015) 'Iconicity in English and Spanish and Its Relation to Lexical Category and Age of Acquisition', *PLoS One*, 10/9: e0137147.

R Core Team. (2020). *R: A language and environment for statistical computing. Vienna, Austria: Foundation for Statistical Computing*.

Ramachandran, V. S., and Hubbard, E. M. (2001) 'Synaesthesia—a Window into Perception, Thought and Language', *Journal of Consciousness Studies*, 8/12: 3–34

Rendall, D. et al. (2005) 'Pitch (F0) and Formant Profiles of Human Vowels and Vowel-like Baboon Grunts: The Role of Vocalizer Body Size and Voice-Acoustic Allometry', *The Journal of the Acoustical Society of America*, 117: 944–55.

Rogers, S. K., and Ross, A. S. (1975) 'A Cross-Cultural Test of the Maluma-Takete Phenomenon', *Perception*, 4/1: 105–6.

Sapir, E. (1929) 'A Study in Phonetic Symbolism', *Journal of Experimental Psychology*, 12/3: 225–39.

Sidhu, D. M et al. (2019) 'Does the Name Say It All? Investigating Phoneme-Personality Sound Symbolism in First Names', *Journal of Experimental Psychology: General*, 148: 1595–614

Spike, M. et al. (2017) 'Minimal Requirements for the Emergence of Learned Signaling', *Cognitive Science*, 41/3: 623–58.

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*, 8/4: 2041669517724807.

Swadesh, M. (1971). *The Origin and Diversification of Language*. Piscataway, NJ: Transaction Publishers.

Tamariz, M. et al. (2017) 'The Interactive Origin of Iconicity', *Cognitive Science*, 42/1: 334–49.

Taylor, I. K. (1963) 'Phonetic Symbolism Re-Examined', *Psychological Bulletin*, 60/2: 200–9.

———, and Taylor, M. M. (1962) 'Phonetic Symbolism in Four Unrelated Languages', *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, 16/4: 344–56.

Walker, P. et al. (2010) 'Preverbal Infants' Sensitivity to Synaesthetic Cross-Modality Correspondences', *Psychological Science*, 21/1: 21–5.

Westbury, C. et al. (2018) 'Weighing up the Evidence for Sound Symbolism: Distributional Properties Predict Cue Strength', *Journal of Memory and Language*, 99: 122–50.

Wichmann, S., Holman, E. W., and Brown, C. H. (2010) 'Sound Symbolism in Basic Vocabulary', *Entropy*, 12/4: 844–58.

Willners, C. (2001). 'Antonyms in Context: A Corpus-Based Semantic Analysis of Swedish Descriptive Adjectives', Dissertation, Lund: Lund University.

Appendix 1. Phonetic transcriptions of the collected audio recordings

Chain	Generation	CONTROL	SMALL	BIG	POINTY	ROUND
1	0	gimpralhus	gimpralhus	gimpralhus	gimpralhus	gimpralhus
1	1	impralhus	gimbralhus	grimpahus	kimpralhus	gimpralhus
1	2	diŋpraos	gimbrawos	grilpaws	kimpralhu	mumavmuv
1	3	dinkambos	gimbambos	drinpaws	pimpabu	novamo
1	4	dinkaboʃ	gimbapos	drinpaws	pinkavu	nofano
1	5	binkabos	gimbapos	grɛdinpaws	finkawu	nofɑno
1	6	bimkabos	gimbampows	grɛdinfaws	feŋkawu	nofano
1	7	dimkabos	gimbampos	grɛdinfaws	dʒɛŋkaow	afaʁnaw
1	8	difgabos	kitbabos	ɹɛlikbawt	deŋgawo	gasamõ
1	9	disgabu	getpovos	ɹɛlibawʈ	gɛŋgau	gasamo
1	10	diskabu	getpovos	ɹɛlibaʈ	dindau	kasamo
1	11	diskabu	kipevos	ɹɛlibaɹ	dindadu	kazamo
1	12	diskabɹə	tipevos	ɹɛliba	dindaŋdu	hatanmɘɹ
1	13	diskapə	tepivols	ɹɛliba	diŋdaŋdu	hatanmarow
1	14	iskawoɹ	tɛtəfons	tʃʊlifa	dindindu	hatanmajl
1	15	iskawoɹ	kɛrfols	tʃʊdefaj	dendendu	hakanman
2	0	gimpralhus	gimpralhus	gimpralhus	gimpralhus	gimpralhus
2	1	gimpralhus	gimprahus	impralhus	kimpaws	gimprasdu
2	2	gimpalhus	gimprahus	impralhus	kimpows	miprasdu
2	3	minpalfos	hibrahu	inpaws	dɹimpls	mirazdu
2	4	mipapots	sibrahu	inpawʃt	kimpas	miraθdu
2	5	mipafots	sbrahu	inpawʃt	gampas	jumejraθdu
2	6	mipatsfo	sbraku	nuonfawʃt	kɛlfɛʃ	jumejglaθdu
2	7	miashowm	sigolaku	undfawʃt	kalfɛs	jumejdasdu
2	8	miɛsxo	siɲolaku	andfowst	kewsrəs	jumejdasdu
2	9	viɛsxow	siɲolaku	anfowswa	kɹəsəs	tyomedastu
2	10	wihæshow	seɲolako	anfowfa	pɹəsəs	tiomedasdu
2	11	vihæsxo	seɲolako	anfowfa	ɹeses	djumidasdu
2	12	vihɛshow	seɲolako	fofofa	ɹeseses	djumidasdu
2	13	vihɛsho	senolako	hohoha	ɹesesam	piemivastu
2	14	bixɛsho	senolaŋko	hohoho	dɹeseɟɑm	kemivastu
2	15	biçesho	senolanko	xoxoxo	tɹɛseto	tjɛmibasdu
3	0	gimpralhus	gimpralhus	gimpralhus	gimpralhus	gimpralhus
3	1	gintrahos	ginprajhus	grimpahuəs	gimprahos	gimpralhus
3	2	nintrahos	gintɹajhus	pusbansus	gimprahʊs	gikkamus
3	3	mintraxos	intrajhos	tunspanstuns	inpranhos	kamuʃ
3	4	ɛntrahos	intrajhos	daŋspastumf	durangos	ikamoʃ
3	5	ɛntraxos	endɹajdɹos	taməstʊ	durɑŋgos	ikaməʃ
3	6	ɛntrahos	hedɹajdɹas	daɹheɹtʊ	durɛngos	ikamaʃ
3	7	ɛntrahas	hetɹadɹatɹaws	gjaʁɛʁtʊ	dorɛngos	ikafaʃ
3	8	ɛntrahas	ntədʒaɹs	gjaʁɛʁdu	dorɛnkos	ikafas
3	9	hɛndrahos	intendʒas	gjɒʁɛʁdu	vorinkos	ikafas
3	10	tɛndʒoxas	insindʒɑs	jɒʁɛʁdu	horinkas	ikafas
3	11	tɛndʒowhɑn	teɲʃiŋstaŋ	jɒɛrdu	haringas	ikasfas
3	12	tɛndʒowhɑn	tɛɲʃistaŋ	dɒɛrdu	dartingas	iasas
3	13	pɛndʒowxa	tɛɲʃistaŋ	mɒɹendun	gasndas	iasas
3	14	bɛndʒuxam	taɲʃistaŋ	uintu	gasɛmdɑs	iasas
3	15	bɛndʒuxaw	daʃistaw	uintu	gasamdos	iasas

(continued)

(continued)

Chain	Generation	CONTROL	SMALL	BIG	POINTY	ROUND
4	0	gimpralhus	gimpralhus	gimpralhus	gimpralhus	gimpralhus
4	1	gimpralhus	gimpralhus	gimpaħus	gimpxahus	gimprahus
4	2	gimpragjus	iterabʊs	ibriðragost	grirarunts	dimprahus
4	3	ginkavjʊs	entoralyns	iðbibragos	grinarunks	poravos
4	4	ginkafjʊs	empeɹomins	isbiŋagos	grinows	boɖvos
4	5	ginpafjʊs	tempreomins	isbiagos	grinows	porlabos
4	6	istaus	temkromvis	izbiagos	grinowʃ	wohlabo
4	7	tispaus	henkromis	izbiagos	grinowtʃ	bobobo
4	8	tispawt	henskromɛ	ibiagon	krinokʃt	bobobo
4	9	istpawt	hentskromɛ	ibiaɤon	gɹinoʃ	bobobo
4	10	itspawts	henskromɛ	bitsbiapon	ɹieɹnoʃ	bobobo
4	11	itspaks	isklamɛ	itsbiapon	ɹinoliʃ	bobobo
4	12	itspans	pispal	kwitsbiaplon	ɹiŋoliʃ	obobo
4	13	glikwɑns	pispala	plisbiao	linalis	obobo
4	14	gliksɑns	tispara	dɛsbiao	inalis	pobobo
4	15	litwɑn	tispara	bejsbiaow	nales	bowbobow
5	0	gimpralhus	gimpralhus	gimpralhus	gimpralhus	gimpralhus
5	1	gimpralhʊs	gimprahus	gimpralħus	ginpɹahus	gintravos
5	2	impralhos	dimprahus	gimbɽawhos	gɛnpagos	hitravos
5	3	impravasa	dimpraħu	imbrawos	gejŋpagwor	kivtrawow
5	4	imprapasa	dimpraxu	imbrawos	ginbagort	kiprahow
5	5	impropesat	impraxu	imprawas	miabort	kibrahow
5	6	inopropesa	impraho	improbos	miapoɽ	ibɹahow
5	7	inopoposa	pobraho	implobosə	miapol	ibahaw
5	8	iopokosa	kobraho	impabosəm	miapal	ibahaw
5	9	iopokosa	kobraho	impabosa	ɲepaʎ	bivohã
5	10	fipokosopa	kobako	imbabosa	njepal	iwohaw
5	11	fiokoʃoba	kobako	imbaborsa	jʊfo	mohab
5	12	vibokoʃalba	howako	imbabolsa	jʊfo	mohamɛθ
5	13	bɹiʃəboninʃdʒɹ	powako	imbavolsa	jʊwfɹo	mohan
5	14	gɹiʃəboɹnisdan	kowako	imbawozba	dʒofo	moharʊ
5	15	wiʃorajasa	koako	imbuzba	iafom	mahow
6	0	gimpralhus	gimpralhus	gimpralhus	gimpralhus	gimpralhus
6	1	gimpralhus	gindahus	gimprahus	gimpahus	gimpralhus
6	2	gimpralfus	daɹus	ɛmprahos	kikadus	kimpraðhus
6	3	gimpralfus	baɹus	ɛmbahos	pikalun	tʃympralwos
6	4	gimprawfus	baɹuʃ	ɛmbɑkos	likalow	dimbrafos
6	5	endʒawfus	taɹuʃ	ɛmbakos	mekalow	ɛdɔsɔ
6	6	endowfus	taruʃ	ɛmbokos	mɛqχalo	ɛdōsō
6	7	dowdowfus	taruʃ	ɛmbokoʃ	mɛkalo	ædōsō
6	8	dowgofus	karuʃt	ɛmbokoʃ	nɛdalo	ɛdrardaw
6	9	dowkofus	skweniʃ	ɛmbokoʃ	minalo	undratr
6	10	gawgawpus	kweniʃ	ɛmbokoʃ	minalo	ndratr
6	11	gawkapus	kweniʃ	ɛmbokoʃ	minaŋgo	məndratr
6	12	gawkopus	kwejniʃ	ɛmbakoʃ	minlajngwo	mundratr
6	13	gawkompus	kwejniʃ	ɛmbakoʃ	milajgwow	bondabəɹ
6	14	gawkopus	kwejniʃ	hewakoʃ	milajgwo	bodaŋgək
6	15	kowkowplus	wejniʃ	iwəgoʃ	milajwo	undadag
7	0	gimpralhus	gimpralhus	gimpralhus	gimpralhus	gimpralhus
7	1	gimpralhus	gɹiimbɹahahu	grimpralhus	imprashus	gimpɹalɹus
7	2	gimrahut	nɛŋhaahu	grinkraprus	imprasus	infraru
7	3	gimrahu	mɛjaahu	gringrasgrʊs	intraʊs	infraru

(continued)

(continued)

Chain	Generation	CONTROL	SMALL	BIG	POINTY	ROUND
7	4	igaɹʊ	nihahahu	gringranrʊs	intrabus	infagʊw
7	5	igarʊ	hahu	wewiwagʊs	içabow	difaigum
7	6	igaru	jahu	wewiwaŋwʊs	izabo	ifogo
7	7	egaru	jahu	weɽiɽaŋgos	izabo	ifogo
7	8	egarʊ	jahu	weɽihaho	bizabo	ifogo
7	9	egarʊ	jahu	weʁhihaʁho	isawo	ifogo
7	10	ejarʊ	jahʊ	ewibihaʁbiχ	isawo	ifodu
7	11	ejalʊk	bjahu	ivivibixabixa	isawok	infodo
7	12	ejalʊk	mjahu	avirbihabiha	isawok	infodo
7	13	ejalʊk	wiaɹhu	ɛrbirlihæ	pisaloko	inforðo
7	14	ejlu	hiaɹhow	ɛnəmbjurixat	susaloko	okordo
7	15	ejlo	siathaw	inembuɹiha	tusalopo	hokogo
8	0	gimpralhus	gimpralhus	gimpralhus	gimpralhus	gimpralhus
8	1	gimpralkus	gimpɹalhus	gimpɹalhus	imbɹalhʊs	imprahus
8	2	nuŋtʃaguʃ	gejprathus	gimpajlhows	imbrohus	imprahus
8	3	nontʃaguʃ	kejbashus	gimpalows	imbraxus	imprahui
8	4	nowlʃabuʃ	gɹimashʊs	gimpalawʃ	ilbrahus	dimprahui
8	5	munʃambuʃ	gimashus	gimpahawst	ilvaʁuʃ	klowfrabō
8	6	munʃaŋguʃ	jumasows	gimbahawst	dizahuʃ	klofabum
8	7	monʃaguʃ	jumasajs	imbahaws	dizahuʃ	kɹunambɹən
8	8	montʃabuʃ	imosajs	imbahaws	dizahuʃt	tumamboɹndɹ
8	9	montʃabus	ivosajs	ɪmbahaws	disawʃ	toroanbondɛd
8	10	oɹtʃambus	hirosvejs	embahawʃ	jusowʃ	tuandbondi
8	11	ɑɹdʒanpus	hirusves	imbahaws	jusowʃ	tuanbondi
8	12	ɑɹdʒampus	hilesves	ɛmbaws	jusowʃ	duɛnbondi
8	13	ɑɹdʒoŋgus	plisvesves	ɛnbaws	jusowʃ	duɛnhondɹ
8	14	aɹçaŋkus	plisvɛsvɛs	imbows	dʒufows	duebon
8	15	arjugus	dəsəsəs	ɛnbos	dʒufoɹs	duedoŋ
9	0	gimpralhus	gimpralhus	gimpralhus	gimpralhus	gimpralhus
9	1	gimpɹalhus	gimpralhus	gimpralhus	mihabos	gimpralhus
9	2	gimpɹalhus	gimprahus	jʊtraləfʊs	pihabōs	impralhus
9	3	ɲimprawthus	klimplahus	jʊtalefeʃ	piŋkakuʃ	imprahurs
9	4	ɹuɹltamhus	klimplæmpluχ	jutalefejʃ	kuŋkakəʃ	imprabɹos
9	5	kukamkus	plimplæmplux	jukarpiʃ	komkakoʃ	χejbabu
9	6	kukamkus	plimplæmpluxta	jukalpitʃ	koŋkakoʃ	hejbabu
9	7	okago	linlænplutr	jukolpitʃ	kodʒakuʃ	hejbabun
9	8	owkaʔow	limlæmluɽa	jugoltitʃ	ozauʃ	hejbabu
9	9	owkraow	liŋlaŋloka	dʒurbarkitʃ	uʃaudz	sejbabu
9	10	howʔaow	miananoka	ubarkitʃ	iʃakɨz	sejbabu
9	11	howawow	miatnanoka	duparkitʃ	iʃakes	sejbopu
9	12	howəow	mianatnoka	dupoɽkitʃ	iʃekis	sejbupu
9	13	howawow	aploka	duborkitʃ	edʒekis	sejbupu
9	14	owawow	apnuka	duborkitʃ	ejekis	sejbʊpu
9	15	owbaow	apnoka	dubodrɪtʃ	iakis	subuku
10	0	gimpralhus	gimpralhus	gimpralhus	gimpralhus	gimpralhus
10	1	gɹinɹaɹus	gimprahus	gimprahus	kybralgus	gimpralhus
10	2	gɹingɹagɹows	gimprahus	kimbrahus	kimpɹagows	timpalhus
10	3	gɹiwgɹawgɹows	biprasus	kimbrahus	kimprakus	timpalohuws
10	4	ɹiwɹawɹows	bibrasu	kimbahus	dibragtus	pimpalus
10	5	gɹewɹawɹows	livlasu	kiwanɛkuʃ	tepraktus	lɛpalows
10	6	dejoramorow	bliblasu	iwankytʃ	difɹatʃuf	klipalos
10	7	eoramoro	bliblasu	deonkətʃ	dipɹɛsəm	klipalos

(continued)

(continued)

Chain	Generation	CONTROL	SMALL	BIG	POINTY	ROUND
10	8	eoramuo	liplasku	dʒunkutʃ	diplɛsɛm	klitaloz
10	9	euramor	nisbetsku	dʒukɛtʃ	iplɛsən	kikitapos
10	10	mewromoʃ	niʃpaskul	dʒukɛtʃ	piplesm	pikitabos
10	11	meromows	nisbasku	dʒukɛtʃ	bipeflum	pikitabos
10	12	mewuʃ	isbasku	dʒutɛtʃ	bipemflum	biɣitavos
10	13	mamuʃ	isbosku	dʒutɛtʃ	dipeflum	bihitahos
10	14	mamuʃ	isbosku	ʒutɛkʃ	distefɹum	bitovs
10	15	mamuʃ	ejsmoskju	dʒəɾɛʃ	distefron	bitows
11	0	gimpralhus	gimpralhus	gimpralhus	gimpralhus	gimpralhus
11	1	gimpralhus	gimprahus	gimprahus	gimprawngus	gimprahus
11	2	ibrarus	gimprafus	diŋkankrus	gimprambus	gejmprahus
11	3	imbralus	dʒimprahus	viŋkaɲkus	gimprambus	gejmpɹahʊs
11	4	ibraluʃ	dimprəhuskwi	viŋkaɲkus	gimprambəs	dejpɹahows
11	5	ofalaruʃ	dinkrəbɵkwi	gɹinkalkus	gimbrombos	tejprahows
11	6	ɵbaluʃ	likrogwisbi	kalkus	imbrombros	deprahows
11	7	ɵdbalɵʃ	likrikwisprik	talkoks	imbrobro	denpawhaws
11	8	ibɹauʒ	ejpimispɹik	talkots	imbrubrump	tenpawhaws
11	9	ibals	ejbigmistɹik	tajgots	ejngoŋgoŋ	ibajhaw
11	10	ibls	bipimisri	ɹajdots	eŋkoŋkoŋ	ajgok
11	11	ivəs	pepemistɹi	dajdos	eŋkoŋkoŋ	ajkot
11	12	ivəs	pepemistɹi	dajdos	ejnskomkom	hajklo
11	13	ivows	efəɹmestri	ajdnos	dinsonklat	hajklow
11	14	iɹoθ	defaɹmestri	teknas	dejnsonkrod	hajko
11	15	iɹowdz	ləfaɹməsti	ednas	diəsamkrɔnt	ajiŋgo
12	0	gimpralhus	gimpralhus	gimpralhus	gimpralhus	gimpralhus
12	1	gimpralhus	himpagrus	gilprahus	gipɹagus	gimpralhus
12	2	dimprahus	tʃipalpus	kilbragus	gipɹagu	impralhus
12	3	libɹahʊs	sibabɵ	ibragus	dʒipɹagu	inkalɹus
12	4	librahos	sibabu	imbragwʊs	jepɹagu	ejtkawsgɹow
12	5	liəlvadhos	çidado	ənkrəkos	jipɹagu	ejtgatvow
12	6	wiɑɹhaɹhows	sidado	ʊkakos	jepragu	ejtkladlow
12	7	wihahaharhos	siudandon	lʊkakos	dʒepragu	ejklawsowh
12	8	wiahow	siudandon	kukakos	jevrako	ejklawdzo
12	9	wiɑɹow	iswopa	kukakotʃ	bærako	eglodzo
12	10	bihao	pliplopla	kukakotʃ	bærakɔ	eglozo
12	11	bihao	priplopla	kukakotʃ	mejrata	eglozo
12	12	bihanaw	tetʃotʃo	kokokotʃ	mejdzafa	jegotso
12	13	bihanaw	dejtsotsow	kutuklotʃ	mejapda	jolgoson
12	14	bianaw	dejtsotsow	gowtuklotʃ	lejwaka	tʃololsam
12	15	bianaw	deʈoʈow	gorklatʃ	lejwaka	tʃowlowsom
13	0	gimpralhus	gimpralhus	gimpralhus	gimpralhus	gimpralhus
13	1	gimprahus	gimpɹahus	gimpralgus	ginpralhus	lintanħus
13	2	njytrahuʈ	gidahohi	gimpralhus	intrahuʃ	ɹintawxu
13	3	dɵtrahɵt	idaoħi	gemkrawkus	endahuʃ	lintahu
13	4	detahəts	idaoçi	kejmkɹawlus	indawuʃ	intahu
13	5	detɔhot	idahoki	kejnkɹawlus	indəwəʃ	intahu
13	6	itɑhot	itahoki	kejnkrɑwtut	indaratʃ	intaʔu
13	7	bitohoɖ	dipaħoki	kejnkrɑwtur	indaratʃ	intawu
13	8	ehoton	dibahoki	ejgɹspoʃ	ingorɔtʃ	intabu
13	9	jehotɔm	dibabowktej	ejdgarots	ingoratʃ	intavu
13	10	jetaŋ	dibabote	owbiqa	hiŋgonatʃ	intavu
13	11	jɛtoŋ	tibabokce	owbiat	hiŋglnatʃ	intaluf

(continued)

(continued)

Chain	Generation	CONTROL	SMALL	BIG	POINTY	ROUND
13	12	dʒeton	ibabawtʃe	owviætʃ	iŋgəlnatʃ	intaldu
13	13	dʒetso	ibababawtʃe	owliedʒ	iŋgenjas	impalu
13	14	dʒɛtso	ibababawʃe	nowdʒejz	iŋgenjaʃ	imbalu
13	15	ɛtso	pirowbawʃəɹ	nowzdej	iŋgledaʃ	iŋbalo
14	0	gimpralhus	gimpralhus	gimpralhus	gimpralhus	gimpralhus
14	1	ukahu	impralhust	giŋbawŋhus	gimpralhus	imprahus
14	2	ukahu	kipralhol	impagus	gimfrahus	embraws
14	3	ukakuk	kibraho	limpagus	dimfrahuʃ	albɹus
14	4	oŋkago	kibraho	mentagys	juswawus	ambɹʊs
14	5	okago	kibraho	mentəkes	jusowus	ambrus
14	6	okago	rihahow	bentʃəkes	jusowus	ambrus
14	7	ɹowkarow	ixaho	mantʃəke	dʒusowbus	ambres
14	8	bɹʉkarok	kjiaxho	najntʃəkje	dʒusowwus	ambres
14	9	blʉkaɹok	tiaxno	lantəte	jusowwul	ampres
14	10	biukaɹok	diaknaw	lantete	jusowbl	ampres
14	11	biukaɹak	diaknow	lanteteʃ	jusowbu	ambres
14	12	ʉkaɹɑ	diagnow	lanɹeɹetʃ	pjusowkow	ɑmbres
14	13	ʉkaɹɑ	nianow	lajnməməs	bisongoln	ɔmles
14	14	ukara	leanlʉ	ləwz	bisongow	ɑmplʊs
14	15	ukara	lejawlʉ	lows	imsomgɔ	okulus
15	0	gimpralhus	gimpralhus	gimpralhus	gimpralhus	gimpralhus
15	1	dintahus	gimpralhus	ɹimbɹaho	dʒilpralbus	impralhus
15	2	inpahun	hunpɑnhus	gɹinaho	dʒirprawus	imprambus
15	3	iŋtao	heŋtʃaŋgeʃ	bimaho	dʒypraws	imprambo
15	4	indao	ɛntomnes	mimaho	dufraw	iŋkrambol
15	5	birao	ɛntomnets	mimaho	notkal	mkrambl
15	6	biʔaʔo	ɛnfunet	maho	matkɑt	kapu
15	7	miʔaʔow	enfane	mahowl	tskold	kapu
15	8	miʔaʔo	dinfane	maho	tskot	kapu
15	9	miʔaʔow	insoneŋ	maho	skowt	kapuçi
15	10	miʔaʔo	mensomej	mahow	skɔwt	kapuħi
15	11	miʔaʔo	mensomen	maho	kow	kapuʃi
15	12	miʔaʔo	lɛnʃomæ	bahow	kul	kapuʃi
15	13	miʔaʔo	glenʃomæ	bahaw	kul	apuʃi
15	14	liʔaʔo	kenʃomeɹ	bahaw	kurv	apuʃi
15	15	liʔaʔow	kentʃmɛ	ahaw	kuɹ	hafoksi
16	0	gimpralhus	gimpralhus	gimpralhus	gimpralhus	gimpralhus
16	1	giŋpawwʉs	imprabus	itbawjuʃi	inprahus	gintraŋgus
16	2	gimpawɹʉs	grimpravus	bigmajuʃi	inprafus	mintaŋgows
16	3	impaus	grindabu	igɲuʃiʃi	implafus	nitango
16	4	impæls	etavjo	mijuʃiʃi	emplawse	ɹekotkom
16	5	impæls	edavij	wiuʃiʃi	plawse	ɹekokom
16	6	intal	edabjʊ	weɹoʃiʃi	plawse	xekokom
16	7	inkɑl	dindadjʊ	wejomʃiʃi	pawse	ekokow
16	8	iŋkɑl	diŋdaŋdu	bworiʃiʃi	awsen	ekokow
16	9	iŋkow	iŋgaŋgoŋ	goɾiʃiʃi	awsn	kokoa
16	10	iŋkomut	miŋgaŋgoŋg	goriʃiʃi	ɔlsn	boko
16	11	mejkow	miŋgaŋgoŋ	goriʃiʃi	koltsom	koko
16	12	mejkol	piŋkaŋgo	goriʃiʃi	koltsom	koko
16	13	mejkos	iŋkaŋgoŋ	boiʃiʃi	kowltiŋ	kɔkɔ
16	14	tsmejkaas	iŋkaŋgow	bojʃiki	hawedin	kokow
16	15	listɛnkaɹ	iŋkambo	utʃihi	hawedi	kowkow

(continued)

(continued)

Chain	Generation	CONTROL	SMALL	BIG	POINTY	ROUND
17	0	gimpralhus	gimpralhus	gimpralhus	gimpralhus	gimpralhus
17	1	ginpahus	intrahus	grimpragus	gimprahus	gimprahus
17	2	kiŋgahu	ejtrax	grimpragus	dʒiŋʔaʔus	dʒukɹahus
17	3	tʃiŋgahu	tra	grimbragos	dʒipaus	jokrahus
17	4	diŋgahu	ɹa	gimbragos	tʃiaus	jogabu
17	5	biŋgəhu	ɹaɹ	imbragus	sialgos	dodamu
17	6	biŋgau	fɹaɹ	kumbrakus	sialɣo	dawdanʉ
17	7	bindaw	rar	anrakutʃ	sialgo	dondav
17	8	bindaw	rak	anrakus	sialgo	doŋdaŋ
17	9	lindaw	rak	amɹakəs	sialgo	doŋdoŋ
17	10	lindaw	rak	anɹakos	sialgo	doŋdoŋ
17	11	iŋdaw	raks	anɹækas	sialgo	nambə
17	12	inda	raf	anwætast	sialgo	mambe
17	13	inda	rah	anwættas	tsialgo	mambeɹd
17	14	inda	ra	anwettas	dialgo	mambe
17	15	inda	ɹɑ	armis	dialgo	mambe
18	0	gimpralhus	gimpralhus	gimpralhus	gimpralhus	gimpralhus
18	1	gimpralgus	gimbrahus	gimpɹælhus	gimpralhus	imprahus
18	2	kilfrɑʊs	vinblahus	impræbus	gimpʁahu	infraʔus
18	3	kifɹomwes	çinbləvuʃ	impravus	belkɑl	instraʔuʃ
18	4	dɹifɹamwes	ʃiblahʊs	impahus	belkɑ	mistraʔus
18	5	dɹifɹamɹejs	ʃinaʔus	umpakus	verʔɑ	mistraʔus
18	6	misombɹe	ʃinaʔɔs	uŋkahus	dɛpa	imistraʔus
18	7	isamwe	etʃimaʔɑ	luŋkahun	debɑm	emestrahus
18	8	nawej	zecimaʔaft	luŋkao	dejbo	imestratsus
18	9	nokwe	imaʔaft	loŋtao	dejbow	inedatʃus
18	10	notwa	imaʔaf	loŋtəaw	tejbo	medatʃys
18	11	mopwa	deamaʔaf	lontaŋɑ	dɛjbol	mitatʃʉs
18	12	mopwap	deɹmaʔaf	boŋtaʔɔ	djejbo	metaʃuʃ
18	13	motɑ	denaʔas	ondaʔo	ipo	metadʒəs
18	14	bobo	benaʔas	oɲagow	ipo	metaʒəʃ
18	15	bowow	enaʔas	owlaʔow	iow	kopagjaʃ
19	0	gimpralhus	gimpralhus	gimpralhus	gimpralhus	gimpralhus
19	1	gimpɹalhus	gimprahus	impajprus	gimpɹalhus	gimprabus
19	2	giŋfawsun	dʒimkrokus	viniŋpajkas	enjɔhus	giɹwaɹdus
19	3	njufowntʃy	dʒinkrokus	tiniŋbajpɑs	dedʒawu	dʉbadʉs
19	4	irfontu	dʒimkropus	tininwajpas	dedʒabu	dʉbadʉs
19	5	diontu	dinkrokos	miwajfas	dejabu	dɹubardus
19	6	dionto	tiŋtloŋsʉn	mijʊdefas	bejabu	ufagət
19	7	dʒohomtom	diŋgonstriəm	mijʊefas	plejabu	ofagoɹ
19	8	dʒohomtom	kitjænsɹ	mijʊɹefa	plejabu	ofagɹ
19	9	dʒohontoŋ	tʉkasəɹ	mindʒupdʒa	vejabu	owfagoɹ
19	10	johonto	dintɑstəɹ	meɹdeɹdudax	mejabəw	owbagloɹ
19	11	diotonkɹoŋ	dinkɑstəɹ	deɹdeɹdudɑħ	bək	owbaglowɹ
19	12	biotonkɹoŋ	dikowstəɹ	dʒedʒedʒədʒa	lək	obagloɹ
19	13	fiontonkrom	dirkoster	dʒədʒədʒədʒa	ljuk	obaglor
19	14	fiontonkron	biekoste	tadʒaridʒa	ljuk	wobɹagloɹ
19	15	bijontonkɹon	bikosta	tɹadadidʒa	ljuk	wobræglor
20	0	gimpralhus	gimpralhus	gimpralhus	gimpralhus	gimpralhus
20	1	gimpralhuf	gimpralhus	grimbrahus	fwahows	gimprahus
20	2	binvahuf	gɹifajəs	grimkardhus	fratʃe	hilgasows
20	3	diŋvahuf	pipajəs	primpagus	fɹatse	hiodasus

(continued)

(continued)

Chain	Generation	CONTROL	SMALL	BIG	POINTY	ROUND
20	4	tinvahuf	vifajeɹs	impagwə	gɹatse	biodaskuls
20	5	huhashʉ	vitfajləɹs	imbagwo	pɹətsik	biodasku
20	6	huhæsdu	skɹajn	imawo	hetʃiks	biodasku
20	7	huhasdu	skrojn	ibawof	hetʃik	iedasʔu
20	8	ħuħastu	skəjn	bigbabos	hetʃi	jedashu
20	9	huhasdu	sklejn	bigbagbogbog	poptʃi	jenasplus
20	10	huhɛstu	klejn	igbakokoa	boptʃi	jenasblus
20	11	huhestu	kwejn	idapitow	moktʃi	enwadslost
20	12	uesdu	pwejt	kintaŋkintoŋ	moʔtʃi	dawarlos
20	13	wɛsdu	pleʈ	kintawkintow	moptʃi	imbaluaɹdos
20	14	esdu	plej	diŋdoŋg	moptsi	iŋgbaluargdos
20	15	distu	klej	diŋdoŋ	mɔptʃi	imbalwardos

Appendix 2. Model outputs for proportions of included sound parameters in the control-small, big-, pointy-, and round-conditions

Sound parameter	Condition	Fit	Lower	Upper
HIGH	CONTROL	0.230618656699477	−1.44600765240503	5.89810125366219
HIGH	SMALL	0.215307860182364	−3.59960064200368	5.27781288664448
HIGH	BIG	−0.429324801801926	−7.44853425963917	3.6145469850836
HIGH	POINTY	0.160035347019068	−3.71896312624912	5.14292958106367
HIGH	ROUND	0.205586858922572	−3.73886919920584	5.30514048194895
FRONT	CONTROL	−7.10997139459894	−16.078154127322	−0.0653874815210604
FRONT	SMALL	−13.6995009671562	−21.5792750892134	−6.04640779873769
FRONT	BIG	−3.94734923979815	−11.3975021214564	3.14690034162625
FRONT	POINTY	5.63033111236483	−1.94792308187174	12.4629525911913
FRONT	ROUND	−13.1962810641052	−20.7258959944684	−5.07764807620099
ROUNDED	CONTROL	2.10412984861119	−2.68700911516405	9.45676589246532
ROUNDED	SMALL	12.1885376922047	3.64115587393778	19.8238946356357
ROUNDED	BIG	0.0289810570395517	−7.01511685936072	7.72157972551947
ROUNDED	POINTY	−10.2964476802425	−16.9215588987712	−2.79524826579275
ROUNDED	ROUND	7.41252866818995	−0.336731530599999	15.2298722431993
GRAVE	CONTROL	−15.756488222744	−−21.4013542377785	−10.5959133095439
GRAVE	SMALL	−8.07137606563015	−14.409759352628	−0.840858785745413
GRAVE	BIG	−15.0566313395363	−21.6081114724104	−9.25149273705352
GRAVE	POINTY	−14.0017662877704	−19.3259909710165	−8.2927395205946
GRAVE	ROUND	−7.10762246398289	−13.9379216570631	−0.670727152259605
VOICED	CONTROL	4.77359016450059	−0.0897539246172075	9.96759058046257
VOICED	SMALL	7.25511072671258	1.79133750356536	13.0132111813632
VOICED	BIG	4.03273103477755	−0.921432109624336	9.17931598723252
VOICED	POINTY	−4.44167880747909	−10.4039163885981	1.47397503546779
VOICED	ROUND	2.69896615471487	−2.8808916156315	8.09826945459457
SONORANT	CONTROL	−0.0835566394657477	−3.53409563787351	2.55174344512217
SONORANT	SMALL	−1.40536443840184	−7.29131826788958	1.70071455239954
SONORANT	BIG	−5.64054278646782	−11.7783349076275	0.0807908235741937
SONORANT	POINTY	−0.306888698644833	−4.6831638217908	2.9777086059265
SONORANT	ROUND	−0.266274423544893	−4.71102686489531	3.2528852889958

Appendix 3. Model outputs for proportions of included sound parameters in the small, big-, pointy-, and round-conditions compared with the control-condition

Sound parameter	Condition	Fit	Lower	Upper
HIGH	SMALL	−0.0158372121688544	−5.22494111352228	4.80672305376131
HIGH	BIG	−0.808793356097368	−8.62398201450861	3.084707554482
HIGH	POINTY	−0.00607306425445842	−4.88362484590099	5.06056663036227
HIGH	ROUND	0.0302836527710895	−2.67591384161046	5.45072363299348
FRONT	SMALL	18.7527168640107	8.30817859370361	27.8608111894988
FRONT	BIG	9.29950360253686	−1.18085739527952	18.58767658783
FRONT	POINTY	−0.578850725356183	−11.1495441929333	8.79977776628002
FRONT	ROUND	5.73914863972107	−1.99952472789698	15.1919997896965
ROUNDED	SMALL	−17.7847631647499	−26.963339736036	−7.43965971186916
ROUNDED	BIG	−7.40349066298248	−17.2297291450552	2.78594957374314
ROUNDED	POINTY	4.60534902124532	−5.66039084782107	14.8646268711208
ROUNDED	ROUND	−4.79664394458426	−13.6094126592691	2.95395728876415
GRAVE	SMALL	−6.96966978136323	−14.2891073266572	1.47463718748436
GRAVE	BIG	−7.8903248205153	−17.1401146467185	0.166798333208379
GRAVE	POINTY	−0.996415805129882	−9.21245352774291	8.75801669971933
GRAVE	ROUND	−8.69270568557871	−16.6309591029071	−0.53777084329238
VOICED	SMALL	−7.17571839521653	−14.1185889841863	0.159392159013844
VOICED	BIG	1.11855091906862	−4.53455577210316	7.74855238912453
VOICED	POINTY	4.46756899485332	−1.26613271204654	11.4818834393862
VOICED	ROUND	1.50763964798706	−2.18048088088113	8.50508504208834
SONORANT	SMALL	−0.0192545842130514	−4.5341276333049	4.27287968514856
SONORANT	BIG	−5.105602559082	−11.888138251758	0.630753049861079
SONORANT	POINTY	−0.863689646838814	−7.18329608668868	2.96510172372569
SONORANT	ROUND	0.0171630209289866	−3.30097130426665	4.45269176299176