SOUND SYMBOLISM IN SIZE PERCEPTION: EFFECTS OF CHILD-DIRECTED SPEECH AND STRATEGY CONGRUENCE

László Kovács[a], Renáta Németh[a], Hilke Elsen[b]
[a]*Faculty of Social Sciences, Eötvös Loránd University, HU*
[b]*Ludwig-Maximilians-University Munich, DE*
kovacs.laszlo@sek.elte.hu, nemeth.renata@tatk.elte.hu, hilkee@lrz.uni-muenchen.de

Abstract. Sound symbolism – the systematic association between phonetic properties and meaning – plays a prominent role in size ratings, where particular speech sounds shape perceptions of magnitude. This paper explores two previously understudied dimensions of size-sound symbolism: (1) the impact of child-directed speech (CDS) on size judgments, and (2) the alignment between participants' self-reported and actual rating strategies. The first section provides a concise overview of the theoretical background on sound symbolism and its relevance to size perception. The second section presents new analyses of an existing dataset, drawing on empirical data from Hungarian and German speakers. Our findings reveal a robust diminutive effect: artificial words that resemble CDS forms are consistently rated as smaller, even after accounting for other phonetic variables. Furthermore, we find that at least one explicitly named strategy – vowel-based judgments – closely mirrors the strategies participants actually employ.

Keywords: sound symbolism, size-sound symbolism, size rating, child-directed speech

DOI: https://doi.org/10.12697/jeful.2026.17.2.02

1. Introduction

Sound symbolism describes the phenomenon when sounds of words bear a non-arbitrari relationship to their meanings. Sound symbolic effects are analyzed and discussed in psychology (e.g., Sidhu & Pexman 2018), in literature (e.g., Elsen 2016; Ferber 2019; Fónagy 1961; Tsur 1992 2012) and in linguistics (e.g., Benczes 2018; Elsen 2016; Hinton, Nichols & Ohala 1994; Lockwood & Dingemanse 2015). Research on sound symbolism in linguistics includes the description of invented languages (Elsen 2016, 2018, 2019), the analysis of rhymes (Benczes 2018), typological patterns (Blasi et al. 2016), language and word learning (D'Anselmo et al. 2019; Nielsen & Dingemanse 2021) as well as word processing (Sidhu, Viglicco & Pexman 2019). It has also been extended to new (sub)fields like the analysis of grammar (Nuckolls 1996), colors (Erben Johansson, Anikin & Aseyev 2019), sign language (Thompson et al. 2020), bodily movement (Shinohara, Kawahara & Tanaka 2020) brand names (Jun & Lee 2022; Klink & Wu 2014; Shrum et al. 2012, Wu, Klink & Guo 2013;), research on first language acquisition (Imai & Kita 2014; Imai et al. 2008; Kantartzis 2011; Laing 2019), the analysis of invented names such as Pokémon names, and names for fantastic creatures (Elsen 2016, 2017, 2018; Kawahara 2020; Kawahara & Breiss 2021; Kawahara, Godoy & Kumagai 2020; Kawahara & Kumagai 2019; Kawahara, Noto & Kumagai 2018) or to different languages of the same language family (cf. e.g., Heinsoo & Saar 2015).

Several studies (for example, Blasi et al. 2016; Ćwiek et al. 2024; Erben Johansson et al. 2020; Joo 2020) suggest the language-independent character of at least some phenomena connected to sound symbolism.

An extensively researched type of sound symbolism is size sound symbolism (cf. Ekström 2022; Winter 2025; Winter & Perlman 2021). Size sound symbolism assumes that sounds are non-arbitrarily associated with the size of objects (cf. Ekström 2022). Research on the topic has shown that in several languages, back vowels are associated with a larger size, while front vowels with a smaller size (cf. Elsen, Németh & Kovács 2021; Knoeferle et al. 2017; Shinohara & Kawahara 2010). Research, however, failed to show a similar general effect of consonants: some research could suggest, for example, that voiced consonants seem to evoke larger size compared to voiceless ones (cf. Saji et al. 2019 for Japanese), but results are not generalizable, since findings on a language are often not found in another language (Ekström 2022). Diffloth (1994), however, demonstrated that identical acoustic contrasts can correspond to opposite size meanings across languages, depending on the prevailing cultural ideology. This observation highlights that sound–meaning associations are not universally fixed but culturally mediated. Contrasts may thus convey opposing size associations because different cultural frameworks map meanings onto the same acoustic poles in divergent ways.

Sound symbolic devices also influence how small children perceive objects. Research has shown that four-month-olds associate the vowel /i/with smaller objects and the vowel /a/with larger ones (Peña, Mehler & Nespor 2011). Research has also shown that children are sensitive to Köhler's (1929) often cited "bouba-kiki" effect and categorize artificial words as spiky or round (Pejovic & Molnar 2017). Imai and Kita (2014) also elaborate their "sound symbolism bootstrapping hypothesis", in which they assume that sound symbolic forms function as a kind of reference. We must separate, however, vowel-based size symbolism from f0-based frequency-code symbolism: lower f0 (lower pitch) is linked to larger speaker body size and dominance, while higher f0 signals smaller speaker body size and submission. The effect, however, does not transfer to the referent objects (Ćwiek & Fuchs 2025). These associations are rooted in cross-species bioacoustic patterns and systematically influence how listeners interpret size in speech (Pisanski et al. 2017).

In another study, Laing et al. (2025) could not show that size-sound symbolism (more exactly, pitch-size congruency) often occurs in child-directed speech in English. Numerous empirical studies have demonstrated, however, that English child-directed speech contains a substantial degree of iconicity, particularly in the use of sound-symbolic and onomatopoeic words, which can facilitate early vocabulary acquisition (Campisi & Özyürek 2013; Han, Nie & Gu 2024; Perry et al. 2018, 2021; de Ruiter et al. 2018).

Cross-linguistic evidence robustly supports sound symbolism, with studies confirming its effects across diverse languages (e.g., English, Japanese, Spanish, Basque), age groups (infants to adults), and methodologies (EEG, eye-tracking, NIRS) (Laing et al. 2025). Meta-analyses also reveal a moderate, age-sensitive influence on shape-word mapping, consistent across languages with and without rich sound-symbolic lexicons, indicating universal sensitivity (Laing et al. 2025). Moreover, cross-linguistic patterns relating to size and emotions—observed even in children—indicate that sound symbolism may play a significant role in language development (see Kantartzis 2011).

In most cultures, adults modify their language style when talking to children. This style or register is labeled by various terms such as child-directed speech (CDS), infant-directed speech, motherese, parentese, caretaker talk, caregiver speech, nursery talk, or baby talk (cf. Saxton

2009). CDS is a specific instance and register of language use by parents and others when speaking to babies and young children (cf. Foulkes, Docherty & Watt 2005; Saxton 2009). Adults, and often older children as well, tend to speak more slowly, more accurately, and more clearly, with repetitions and pauses, shorter utterances, simpler constructions, and a higher and more variable pitch when conversing with children (cf. Saxton 2009). The vocabulary is simple and covers words for things and concepts in the immediate situation and the child's interests. This style exhibits both affective and simplifying qualities. Vowels are especially hyperarticulated (Burnham, Kitamura & Vollmer-Conna 2002). Exaggerated intonational characteristics capture the children's attention, and help them focus on communication, and thereby on social interaction (Cooper & Aslin 1990). The role and significance of child-directed speech (CDS) in the language acquisition process have been widely recognized in the literature (Golinkoff et al. 2015; Outters et al. 2020). One of CDS's characteristics is the use of words that exhibit simpler sound patterns than those of adult words, such as reduplicated syllables, as seen in words like *choo-choo*, *dadda*, and *num-num* (cf. Garmann et al. 2019).

Reduplication, the repetition of syllables or phonological units within a word, plays a significant role in various aspects of language acquisition and processing. Research by Ota and Skarabela (2018) demonstrates that reduplication facilitates segmentation, helping infants better parse continuous speech into meaningful units, and that it also influences vocabulary development in children (Ota & Skarabela 2016). Additionally, reduplication of consonants influences adults learning English as a second language (Basnak & Ota 2024). Similarly, recent research by Cathcart (2024) shows that consonant and vowel repetition within reduplication yield distinct cognitive and perceptual effects, suggesting that the type of phonemic repetition modulates how reduplicated forms are processed. This distinction signifies that the structural composition of reduplication influences its function in language acquisition and processing, with consonant and vowel repetition engaging distinct cognitive mechanisms.

Studies also show that reduplication intensifies evaluative meanings such as smallness and affection, functioning as a salient cue in iconic sound symbolism (Akinbo 2023), and that reduplication expresses smallness or diminutiveness (Kouwenberg & LaCharité 2005). Motoki, Iseki & Pathak (2025) hypothesized that reduplication triggers associations with the baby-schema cuteness.

In our study, some of the stimuli we use resemble such children's or family terms, e.g., German *baba* 'dirty',[1] *Pipi* 'urine', *Popo* 'buttocks', *Kaka* 'excrement' and *dada* (*dada gehen*, 'to have a stroll'). We seek an answer to the question of whether artificial words resembling those from child-directed speech are more likely to be characterized as very small/small objects, even when their constituent sounds predict a different size.

Sometimes experimenters asked participants to explain their decisions. Asking for reasons and one's own reflection on empirical data connected to sound symbolism is not new in research. As early as 1924, Usnadze presented six "senseless" drawings to ten test participants. They had to choose the most appropriate made-up word for the figure from a list of made-up words and explain their decision. In most cases participants created some mental picture: for example, wavy lines were imagined to be tongues of flame. The influence of their mother tongue was shown in part by the fact that the participants tried to name the object mentally. The word they thought of then influenced their choice of the made-up word that was acoustically similar to the intended one. In some cases, the artificial words also sounded like a property that the participant associated with the drawing.

Wissemann (1954) carried out tests with German-speaking participants to investigate the creation of new terms for sounds. They had to listen to 14 different noises or noise complexes, for example, hammering against an iron weight, or splashing with water. In one experiment, participants had to choose from a list of artificial words the one that they thought best suited to name the noise they heard; in another, they had to create suitable artificial words themselves. Self-observations of the participants revealed that noises are never purely acoustic shapes, as both linguistic and non-linguistic aspects of the noise phenomenon influence perception. For example, when listening to splashing, participants perceived water even though they did not see it. Several made-up words were created by varying existing words having to do with the perceived phenomenon, like *zischen* 'to hiss': *ptzschischsch; plätschern* 'to

1 In Austria this is used to say goodbye. This interpretation was neglected, since the data was not collected in Austria.

splash' (Wissemann 1954: 67f.). The influence of their mother tongue, and the linguistic environment of the noise became apparent.

In Elsen's (2008) study, German participants were given six names from fantasy and science fiction stories to choose from as more or less suitable for certain creatures, such as magicians, ugly, evil monsters, or small, harmless beings. The names had to be rated on a scale between very suitable (1) and not suitable at all (7). One of the tasks was "How well does the name suit a beautiful young woman? (*Klipp*, *Alani*, *Olda*, *Valeron Veit*, *Sartassa*, *Gorx*)." In the following discussion, participants were asked about their criteria for rating. Some decisions were based on sound: light vowels were important for distinguishing women's names, while final consonants clearly indicated male characters. A further factor was the number of syllables, combined with the sounds. In the case of the evil monsters, throat sounds were clear indications of the language that the beings were thought to speak; this was especially true for *Ch'tuon* and *Chrekt-Orn*, a connection between velar and uvular consonants and danger is found in several studies (Elsen 2016). Owing to their shortness, *Klipp* and *Gorx* sounded hard, sharp, angular, and therefore clearly not feminine. Overall, the test participants rated names with consonant groups and few vowels as hard. *Krillri* and *Cir* sounded like a bright, high-pitched voice and were therefore a good fit for dwarves, elves, and other graceful creatures.

Taken together, participants' self-observations reveal that one of the reasons for their decisions was sound symbolism (*Tik, Cir* – /i/: light, bright, small, harmless). However, similarity to existing words having to do with associations around the object or creature was another reason (cf. *Sartassa – rassig*, thus beautiful, thus a suitable name for a beautiful young woman; *Veit* like *Veitstanz*, thus negative). Overall results showed that the influence of existing words on sound rating can never be completely ruled out.

1.1. Research questions and hypotheses

1.1.1. Research Question 1

Taking into account the sound symbolism bootstrapping hypothesis of Imai and Kita (2014), the question arises whether learned sound symbolic associations, for example, acquired while exposed to child-directed speech, may also influence size ratings.

Although from the analyzed artificial words, none of the three repetitive syllables from the analyzed artificial words form a Hungarian or German word, some two-syllable versions are, however, an element of child-directed speech (CDS) in both languages. The question arises, therefore, whether the three-syllable stimuli with resemblance to CDS-words were evaluated differently by the participants, the hypothesis being that if such an effect occurred, it was to reduce the perceived size of the marked object through association with babies. CDS makes use of hypocoristics, among other strategies, and since smallness and hypocoristic meanings often go hand in hand, for example, in German and in the case of English diminutives (Kempe, Brooks & Gillis 2005), we assume that the three-syllable versions might inherit these connotations.

To test for this possibility, we analyzed how size ratings of artificial words resembling words of child-directed speech affect the size rating of objects. More specifically: were stimuli resembling words from child-directed speech more likely characterized as very small/small objects even if their constituting sounds would predict another size? Consequently, the influence of the perceived meaning of the stimuli (i.e., the resemblance of the stimuli to existing words) is addressed by the first research question.

RQ1: Do the resemblance of artificial words to words from child-directed speech affect participants' ratings?

H1: Words that resemble CDS forms are consistently rated as smaller in size, even when other phonetic variables are taken into account.

1.1.2. Research Question 2

The results and considerations of the above research show that analyzing the strategies connected to sound symbolic decisions may lead to new insights. Thus, another key question in size rating within sound symbolism research is whether the actual strategies participants use to judge size are the same as the explicit, named strategies they report using. This addresses the potential gap between participants' self-reported reasoning and the unconscious or automatic processes that may actually drive their size judgments.

In our study, we also asked participants to name their strategies upon which they based their size ratings for the objects. With RQ2, we aimed to determine whether the participants' stated – conscious or unconscious – strategies were indeed the ones they used to rate the size of the objects.

RQ2A: What conscious strategies are used by the participants in their sound-size ratings;

RQ2B: are these consciously named strategies congruent with the actual strategies used?

H2A: Since linguistics students did not participate in the survey, we did not expect reported strategies based on subtle phonetic differences but rather assumed that most strategies would involve simple references to vowels or consonants.

H2B: The reported strategies closely reflect the unconscious strategies actually used by the participants. Based on H2A, we specifically assume that the effect of vowel features is stronger among those who reported using vowel-based conscious strategies.

2. Methods

In our data collection, the initial aim was to compare influencing factors on size ratings in two unrelated languages, namely German and Hungarian, an Indo-European, Germanic language and a Uralic, Ugric language, respectively. The collected data was used in earlier papers to show that size rating is influenced not only by voiced consonants and back vowels, but also lip rounding (Elsen, Németh & Kovács 2021) and that although the characteristics of participants do not affect overall size ratings, they may modify (intensify or weaken) the effect of phonetic features on size ratings (Kovács, Németh & Elsen 2024). The current study offers a re-examination of previously gathered data through a novel analytical approach connected to CDS and stated size-rating strategies.

To facilitate direct comparison with earlier findings, we followed the experimental procedure introduced by Shinohara and Kawahara

(2010) and, where feasible, employed the same set of phonemes. We likewise adopted their analytical approach, using mixed-effects models to disentangle participant- and stimulus-level variation. This parallel methodology allows our results to be meaningfully compared with their findings for Japanese, Chinese, Korean, and English. We have made similar decisions in our previously published series of studies (Elsen, Németh & Kovács 2021; Kovács, Németh & Elsen 2024), so our current results are comparable with those (see, for example, the comparability of the effects of strategies or CDS words with the effect of the gender of the test subjects). Regarding this decision, see also our comments mentioned in the limitations section.

2.1. Material

The data collection is based on the research of Shinohara and Kawahara (2010). They tested native speakers of English, Chinese, Japanese and Korean with 40 artificial words to identify the factors influencing size ratings. The artificial words were doubled VC-syllables, containing /b, d, g, z – p, t, k, s/ and /i, u, e, o, a/ (for example, Shinohara & Kawahara 2010). According to their results, the height of vowels, the backness of vowels, and voicing in obstruents influence size ratings.

The experiment for current analysis is a modified version of Shinohara and Kawahara's research design. Instead of double VC-syllables we used CVCVCV structures. The consonants and vowels we used were three voiced stops /b, d, g/, one voiced fricative /v/, three voiceless stops /p, t, k/, one voiceless fricative /f/ and we used seven vowels /i, u, e, o, a/ and /ø/ and /y/. The last two vowels are vowels that exist in the two analyzed languages, German and Hungarian. Consequently, the artificial words contained seven vowels: back /a, o, u/ versus front /i, y, e, ø/ and high /i, y, u/ versus middle /e, o, ø/ versus low /a/, respectively. Instead of VCVC structures, as used by Shinohara and Kawahara (2010) we decided to use CVCVCV structures on one side to avoid resemblance to existing VCVC words, as many bisyllabic VCVC combinations are lexicalized in Hungarian or in German. For example, in Hungarian, forms such as *tata* and *papa* denote 'grandpa/old man', while *baba* means 'baby' and *popo*, *bebe*, *didi*, and *bibi* are widely recognized as words in children's language. Similarly, in German, *Papa* refers to 'father' or 'daddy', while other CVCV words are names (*Pepe*, *Sisi*, *Kiki*) or terms from children's language (*Pipi*, *Popo*, *Kaka*, *Baba*, *Dada*). The other reason for choosing CVCVCV structures was to avoid the effect of final devoicing of voiced obstruents in German (*Auslautverhärtung*).

The three factors (7 vowels, 4 consonants with 2 voicing types) were fully crossed, resulting in 56 artificial words (7*4*2). In the artificial words, the consonants and vowels were identical (e.g., *kükükü*, *pupupu*, *tötötö*). See the full list of stimuli in alphabetical order in Appendix 1.

2.2. Participants

Participants were students from the University of Augsburg (Germany) and Eötvös Loránd University (Hungary). The study used a non-probability sampling method. Data analysis began with a data screening and cleaning phase. Nine participants recruited in Germany were removed because German was neither their native language nor a language in which they reported sufficient proficiency. In addition, responses deemed unreliable were excluded. Potentially dishonest responding was identified using person–total correlation diagnostics (Dupuis, Meier & Cuneo 2019), implemented via the R script provided by Motamedi et al. (2019), which detects response patterns that deviate markedly from the overall sample. Respondents with non-positive person–total correlations were subsequently removed (10 in the German sample and 20 in the Hungarian sample), including cases suggestive of repetitive answering. After these exclusions, the final samples comprised 281 German and 101 Hungarian participants. The exclusion procedure is validated by the fact that the excluded cases included all the cases that had previously been heuristically identified by the researchers as showing an apparent repetitive pattern. In addition, the results are robust to data exclusions; the same variables remained significant in our models, with the same signs, and the magnitude of the effects did not change significantly. Finally, the same procedure was used in a paper investigating the same data (Kovács, Németh & Elsen 2024), so the results of the two papers are comparable.

Table 1 shows the demographic characteristics of the final sample. The median age was 22 years in Hungary and 21 years in Germany. The gender distribution was uneven, with males representing 46% of the Hungarian sample and 19% of the German sample.

Table 1. Demographic characteristics of the sample.

	Sample size	Gender		Age		
		women	men	median	range	standard deviation
Germany	281	231	50	21	(18;49)	4.0
Hungary	101	56	45	22	(19;53)	7.7

2.3. Procedure

To minimize context effects such as fatigue, we used ten randomized versions of the 56-item stimulus list, each with a different order, and applied the same lists at both data collection sites.

The study employed a paper-and-pencil methodology, with stimuli presented to participants in print, on sheets using Arial font, size 12 pt. Stress was neither controlled nor indicated in any way, since in the original experiment it was also not indicated. We decided against indicating stress or providing an auditory presentation, as doing so would have required us to determine stress patterns and pronunciation before the experiment began. Had we made such determinations, it would have been appropriate to follow the phonological rules of each respective language, which would have been undesirable. Therefore, we opted for a neutral solution that avoided language-specific influences, since this setting was also used in the experiment of Shinohara and Kawahara (2010).

In the experiment, participants rated the assumed size (very small, relatively small, relatively large, very large) of the object – a gem – described by the artificial words. We asked each participant to provide additional information, including their mother tongue, the languages they speak (along with self-assessed proficiency for each on a 4-point scale), gender, and age. Participants were also asked to specify any factors that they think influenced their decision.

2.4. Variables and models used in the analysis

Outcome variable: rating of the size of the given word on a 1–4 scale (1=very small, 2=relatively small, 3=relatively large, 4=very large).

Characteristics of the stimuli (predictor variables): vowel backness, vowel height, voicing of the consonant, roundness, and manner of articulation.

Characteristics of the participants: survey site (Germany/Hungary).

In our previous research (Elsen, Németh & Kovács 2021; Kovács, Németh & Elsen 2024) we also used the variables age, gender, mother tongue, other spoken languages with level of knowledge.

Derivation of predictor variables used in the analysis: The variable 'language knowledge' indicates whether participants speak any languages – other than their mother tongue – from the Germanic, Romance, Slavic, or other language families, at any proficiency level. If a participant reported a second native language besides Hungarian or German, this was classified as a foreign language. Nearly 99% of both samples spoke a Germanic language, while only 4% spoke a Slavic language, resulting in highly uneven distributions; therefore, these variables were excluded from further analysis. In contrast, knowledge of a Romance language was more evenly distributed (7% in Hungary, 85% in Germany) and was included as a variable in the analysis. Language categories were determined based on the collected data, with 'other' languages appearing only once or twice in the dataset. All language knowledge was self-reported. We defined a three-category variable that refers to the number of spoken languages (1, 2 or 3+).

Level of language knowledge: The derived variable refers to the level of language knowledge for the second, third, etc., language the participant speaks best, based on the participant's subjective self-classification, which may be influenced by common language knowledge in the participant's country, i.e., on the survey site.[2] The three used categories are: (1) the participant speaks any language fluently other than his/her (first) mother tongue, (2) the level of knowledge is good, (3) the knowledge level is intermediate / basic or no other language is spoken. In the regression models, we used a dichotomized variable (worse than

2 In the analysis, we use the phrases in the countries referring to the site of data collection instead of referring to the language itself. We decided for these phrasings in light of the results of Elsen, Németh and Kovács (2021): since language knowledge influences at least the perceived role of backness, we find it more appropriate to speak in relation to the results of data collection sites rather than of the effects of the individual languages.

good vs. at least good level). Results shows Hungarians classify their own language knowledge much worse than Germans: In Germany, 95.53% of participants report being good or fluent in at least one additional language, whereas in Hungary, only 69.42% of participants state that they speak a second language well or fluently.

In our models we used **age** as a binary variable (18–23 or 24+).

We use a linear mixed-effect model (Rabe-Hesketh & Skrondal 2012; Snijders & Bosker 2012; for linguistic applications see Baayen 2008) with word and participant as random factors. That is, size ratings (1–4) were analyzed as a continuous outcome to ensure comparability with prior studies, since linear mixed-effects models are commonly applied to Likert-type data in psycholinguistics. Furthermore, it is plausible that respondents also identify the scale of 1-4 with numbers. We acknowledge that ordinal or multinomial mixed-effects models constitute an alternative analytical approach. However, such models impose additional assumptions about category thresholds and their interpretation is also more difficult. Importantly, our primary conclusions concern the direction and relative magnitude of effects rather than exact scale-level predictions.

Since words and participants as random factors are not nested, we used a **crossed random effect model**. We have balanced data (equal sample size in each person/item cell), so restricted maximum likelihood (REML) yields unbiased estimates. In line with recommendations by Barr et al. (2013), we initially attempted to fit a maximal random-effects structure by including all design-justified random slopes; however, these models failed to converge (cf. Bates et al. 2015). Because convergence could not be achieved despite adjustments to the optimization procedure, the random-effects structure was progressively simplified by removing random slopes. Importantly, model estimates remained stable across specifications, indicating that this simplification did not materially affect the conclusions. The final analyses therefore, relied on random-intercept-only models, with size rating as the dependent variable. All models specified random intercepts for participants and stimuli. The explanatory variables were the phonetic features of the stimuli, and we fitted separate models for the Hungarian and German samples.

Categorical predictors were dummy-coded so that each level could be directly compared to a reference category, facilitating straightforward interpretation of the fixed-effect coefficients. This approach also ensures comparability with prior studies using the same modeling framework.

We used Stata's xtmixed command. To leverage the use of mixed-effects models, we aimed **to interpret model parameters that have significant interpretive power**, such as intra-class correlation or proportional change in variance. As these measures are rarely discussed in the size symbolism literature (e.g., not discussed in Shinohara & Kawahara 2010, nor Knoeferle et al. 2017), we introduced them in an earlier paper (Elsen, Németh & Kovács 2021). In this earlier paper, we fitted a model with only the phonetic features as predictors. In the following we will refer to this model as Model 1, because we will answer our research questions by extending this model (we do not need the results of this initial model in this research, but they are available in Elsen, Németh & Kovács 2021).

Different extensions of Model 1 were used to treat our research questions. To answer RQ1, we defined a new word-level variable with two categories (words resembling CDS words / words not resembling CDS words) and added this CDS indicator to Model 1 as fixed effect.

When answering RQ2, we added interaction terms to Model 1 that defined two-way interactions between phonetic features and size-rating strategies.

3. Results

3.1. Child-directed speech

To answer the first research question, we added the CDS indicator to Model 1 as a fixed effect, examined its effect (which we expected to be negative), tested its significance, and checked whether its presence altered the values of the other fixed effect parameters. As words belonging to CDS we chose the stimuli resembling the six Hungarian words: *baba* 'baby', *popó* 'buttocks', *didi* 'breast', *bibi* 'wound', *kaka* 'excrement' and *pipi* 'urine', for German: *baba* 'dirty', *Pipi* 'urine', *Popo* 'buttocks', *Kaka* 'excrement' and *dada* (in *dada gehen* 'have a stroll'). We must note, however, that no lists of usable CDS words exist, either in German or in Hungarian.

Given the experimental design's focus on three-syllable words, our CDS predictor is operationalized by concentrating exclusively on reduplicative forms, which constitute a core phonological characteristic of child-directed speech. We retain the overarching CDS framing, however, as participants explicitly identified CDS resemblance as influencing their size ratings.

A possibility would have been to use the MacArthur-Bates Communicative Development Inventory (CDI), however, the database does not contain reduplicative words in German and in Hungarian, some well-known reduplicative words like *kaka* are missing. Therefore, we selected and designated words as CDS words that are predominantly used when talking to babies, drawing on our own native-speaker competence and expertise, as well as that of colleagues. Our coding, therefore, reflects an operational definition of CDS-like reduplicative forms grounded in shared intuitions among native speakers (see also one of the limitations in this regard).

Table 2 shows the German and Hungarian resembling words together with the stimulus from the experiment they show resemblance with. See the correlation between phonetic features and being a CDS word in Appendix 3 (each correlation is negligible or weak (the absolute value is below .3). Our results are presented in Table 3.

Table 2. CDS words and resembling stimuli.

CDS word in German	**Resembling stimulus**	**CDS word in Hungarian**	**Resembling stimulus**
baba	bababa	*baba*	bababa
didi	dididi	*bibi*	bibibi
Kaka	kakaka	*didi*	dididi
Pipi	pipipi	*kaka*	kakaka
Popo	popopo	*pipi*	pipipi
		popó	popopo

Table 3. Estimates of the original model (Model 1) and the model with the CDS indicator added (Model 2). P-values and confidence intervals in the brackets. Fixed parameters significant at the .001, .0.1, .0.5 level are denoted by ***, **, *, respectively. The numbers in the fixed parts of the table indicate how much larger the object was rated on a 1–4 scale influenced by the given phonetic feature. For example, in Germany in Model 1 words with back vowels were rated 0.35 larger (on a 1–4 scale) than those with front vowels.

	Germany		**Hungary**	
	Model 1	Model 2	Model 1	Model 2
Fixed parts (regression coefficients)				
Backness (reference: Front) *Back*	+0.39*** (0.000) (0.28; 0.50)	+0.40*** (0.000) (0.29; 0.51)	+0.26** (0.000) (0.13; 0.39)	+0.30*** (0.000) (0.17; 0.42)
Roundness (reference: Unrounded) *Rounded*	+0.40*** (0.000) (0.29; 0.51)	+0.39*** (0.000) (0.28; 0.49)	+0.36*** (0.000) (0.20; 0.52)	+0.31*** (0.000) (0.16; 0.46)
Height (reference: Low) *Middle*	–0.14 (0.130) (–0.32; 0.04)	–0.18* (0.047) (–0.36; –0.00)	+0.03 (0.753) (–0.13; 0.18)	+0.04 (0.623) (–0.11; 0.18)
Height (reference: Low) *High*	–0.99*** (0.000) (–1.17; –0.81)	–1.02*** (0.000) (–1.21; –0.85)	–0.34*** (0.000) (–0.47; –0.21)	–0.33*** (0.000) (–0.45; –0.20)
Voicing (reference: Voiceless) *Voiced*	+0.08* (0.047) (0.00; 0.17)	+0.08* (0.058) (–0.00; 0.16)	+0.22*** (0.000) (0.12; 0.32)	+0.22*** (0.000) (0.12; 0.31)
Manner of articulation (reference: Fricative) *Stop*	–0.07 (0.176) (–0.16; 0.03)	–0.05 (0.340) (–0.14; 0.05)	–0.17** (0.007) (–0.29; –0.05)	–0.13* (0.026) (–0.24; –0.02)
CDS (reference: No) *Yes*	–	–0.17* (0.037) (–0.33; –0.01)	–	–0.27*** (0.001) (–0.43; –0.11)
Random parts				
PCV_{St}	–	–7%	–	–21%
$N_{stimuli}$	56	56	56	56
$N_{participants}$	281	281	101	101
Log-restricted likelihood	–19,227	–19,227	–7,354	–7,351
Likelihood Ratio test comparing the two nested models	LR chi2(1) = 4.84 (p=0.028)		LR chi2(1) = 10.99 (p=0.001)	

The CDS-effect on size rating is negative (which is plausible) and significant in both participant groups. That is: if the stimulus is similar to a word from child-directed speech, the participants tend to rate the word to be smaller in size, even adjusted to backness, voicing, etc. (i.e., in case of two words with similar phonetic features, the one that reminds of a CDS word tends to get a "smaller" rating). The effect size is quite strong, its strength is comparable to the effect of the strongest phonetic features (!), especially in Hungary. (Here and hereafter, when we write about effect size, we refer to the raw ordinal difference in size ratings.)

By adjusting for CDS, we may explain some of the stimuli-level variances detected previously in Model 1 (cf. Elsen, Németh & Kovács 2021). The success of this explanation can be measured by the proportional change in variance (PCV) at the stimuli level. The corresponding parameter, PCVSt shows a decrease in variance, and this decrease is really large (21%) in Hungary. That is, the CDS indicator explains some of the between-stimuli variance that remained unexplained even after adjusting to phonetic features.

However, our most important result is that the inclusion of the "CDS" variable did not change the sign or magnitude of the effect of the phonetic features. (The difference between the respective coefficients is only a few hundredths on a scale of 1–4.) This suggests that although it was not possible to define stimuli completely free of semantic associations, the validity of the previous research was not weakened.

3.2. Strategies

The second research question is connected to conscious or unconscious strategies in sound rating. At the end of the questionnaire, the participants were asked to name the strategies they thought had influenced their size ratings: "What factors influenced your decision to rate the given adjective as describing small or large objects?" 95.2% of the German students and 88.4% of the Hungarians answered this question.

Since this was an open question, we categorized the answers after data collection. The following main factors influencing size rating strategies that we identified were: 1) vowels, 2) consonants, 3) similar sounding words, 4) other special strategies, 5) general information on sounds, and 6) nothing specific and 7) mixed strategies. Several of these main strategies were split up into subcategories. The identified main- and substrategies are presented in Appendix 2. In the case of No. 4) (other special strategies), we identified the following influencing factors: associations, word length, imagined speaker, pitch, and stress. General information on sounds (category No. 5) included non-categorizable answers like the answer “sounds”. Mixed strategies may include strategies like “Sounds and expressions from known languages” or “pronunciation”. These we called mixed strategies, since any of the previous strategies may play a role, if size rating is actually based on such a strategy. Appendix 2 gives an overview of the identified strategies and substrategies. Most participants named not only one, but 2 or 3 different strategies that had influenced their decisions. In Table 4 we summarize these strategies. In the case the participant named two substrategies, for example, for vowel-connected strategies, we count vowel-connected strategies twice.

Table 4. The main strategies used by the participants for size-sound ratings. Multiple answers were possible.

Categories	German strategies (281 students)	Hungarian strategies (101 students)
Vowel	159	41
Consonant	33	11
Similar sounding words	67	14
Special strategies	50	16
General information to sounds	45	27
Nothing specific	26	6
Mixed strategies	44	7

To see whether the participants in fact followed the strategy they specified, we could test the vowel or consonant strategy (the first two of the above categories) on our data. However, as few people chose a consonant strategy (especially in Hungary: only 12 participants), we will only focus on the vowel strategy in the following. We operationalized the research question by investigating whether the effect of vowel features is stronger for those who chose this strategy.

We created a new binary (yes/no) variable (‘vowel strategy’) that coded whether someone had chosen the vowel strategy. We added interaction terms to Model 1 that defined two-way interactions between vowel features and vowel strategy. Again, we used Bonferroni correction (we had four interaction parameters corresponding to the three vowel features, so we applied a significance level of 0,0125% instead of 5%). Only parameters defined by significant interactions are shown in Table 5.

Table 5. How vowel-strategy modifies the effect of vowel features on rating (regression coefficients). Only the effect of interaction terms significant at the 5% level are presented (with Bonferroni correction applied). P-values and confidence intervals in the brackets. Effects significant at the level of 0.05, 0.01, 0.001 are denoted by *, **, ***, respectively. Numbers in the table indicate how much larger the object was rated on a 1-4 scale because of the influence of the given phonetic feature.

	Germany	**Hungary**
Backness		
Back vs Front (strategy=no)	+0.26*** (0.000) (.14; 0.37)	+0.22** (0.003) (0.08; 0.36)
Back vs Front (strategy=yes)	+0.50*** (0.000) (0.38; 0.61)	+0.32*** (0.000) (0.17; 0.47)
Roundness		
Rounded vs Unrounded (strategy=no)	–	0.30*** (0.001) (0.13; 0.47)
Rounded vs Unrounded (strategy=yes)	–	0.46*** (0.000) (0.27; 0.64)
Height		
High vs Low (strategy=no)	–	–0.28** (0.000) (–0.43; –0.14)
High vs Low (strategy=yes)	–	–0.42*** (0.000) (–0.57; –0.27)
Log-restricted likelihood	–19184.311	–7344.145

The interaction parameter is significant in case of backness in both participant groups, and the effect size is very strong (e.g., the strength of backness is twice as large among those who chose the strategy than among those who do not).

The interaction parameter for height and roundness is significant only in the Hungarian sample. Similarly to backness, the sign of the strategy's effect is such that it amplifies the strength of the feature's effect. (In case of Germany, the interaction parameters are not only non-significant, but their magnitudes are also negligible.)

4. Discussion

4.1. Child-directed speech

The CDS effect on size ratings is negative and significant in both participant groups: stimuli resembling child-directed speech words are rated as smaller, even after controlling for phonetic features, with an effect size comparable to that of the strongest phonetic predictors, particularly in Hungary. Including the CDS variable substantially reduced unexplained between-stimulus variance, indicating that CDS accounts for variance not captured by phonetics alone. Crucially, adding CDS did not alter the direction or magnitude of phonetic effects, supporting the robustness of earlier findings despite unavoidable semantic associations.

Diminutive forms are often used in child-directed speech, where they function not only as simplification, but to express smallness create a sense of familiarity and affection (Laalo & Argus, 2020). This repeated association between diminutives and communicative contexts involving children strengthens the conceptual link between such forms and the notion of smallness. Therefore, linguistic forms that resemble child-directed speech may implicitly inherit this diminutive quality, conveying subtle semantic or affective meanings related to smallness or endearment. That would mean that every word similar to those in child-directed speech is automatically rated as describing small things or used in a context of smallness. This finding aligns with the results of Dressler, Lettner, and Korecky-Kröll (2012), who note that the pragmatic meanings of diminutive forms are acquired in early childhood and play a more significant role than those of words denoting smallness.

This assumption may be connected to the question posed by Sidhu (2025), who considered what kinds of causal directionality exist, thus whether phoneme associations shape the lexicon, or they bias participant ratings. Connected to his considerations, we may assume that in the case of words resembling CDS, the actual characteristics of vowels and consonants may be overridden by the experiences connected to the use of CDS words.

That CDS influences size ratings of words resembling CDS-words was also corroborated by the participants themselves: 13 German and 2 Hungarian participants indicated in response to the strategy-related question that child-directed speech (CDS) words influenced their decision-making during the experiment. They named strategies like "connected to babies", "children's sounds" or "baby language." This shows that any resemblance to CDS (or other existing words) may influence size ratings.

4.2. Strategies

The above-mentioned research questions can be answered as follows: connected to the first part of the 2nd research question "A) What conscious strategies are used by the participants in their sound-size ratings", we showed that the actual strategies participants used were connected both in Hungary and Germany to vowels, consonants, to similar sounding words, and to special strategies.

After the results, we also asked whether the existence of a conscious vowel-strategy, or the specific content of this strategy, influences how the participant actually evaluated the experimental stimuli?

The existence of significant interaction parameters can answer this question. These parameters showed that the existence of the vowel strategy had an impact on the ratings in both participant groups. The strategy modified the effect of backness the most – this modification is significant and consistent in both participant groups (in terms of sign). In Hungary, the strategy also modified the effect of height and roundness.

Connected the second part of the research question "B) Are these consciously named strategies congruent with the actual strategies used?", the sign of the significant interaction parameters answers this question. According to them, the effect of some vowel features is significantly different for those who chose the strategy, and the sign of the difference is such that the presence of the strategy amplifies the effect of the vowel features. That is, our data showed that the existence of a conscious vowel strategy is congruent with the strategies actually used.

Backness showed a significant and robust interaction with strategy in both participant groups, with effects roughly twice as large among strategy users; interactions with height and roundness were significant only in the Hungarian sample and similarly amplified feature effects. In the German sample, interactions for height and roundness were negligible and non-significant.

The specific content of the vowel strategy influenced how participants actually evaluated the experimental stimuli. The robustness of our results is enhanced by the fact that the effect of the strategy is consistent in the sense that if it exists, it always reinforces the effect of the vowel features. According to our results, the conscious vowel strategy was more extensive in the Hungarian sample: Hungarians were more conscious, for example, of the roundness-unroundness contrast. (Note: Vainio 2021, has also shown that vowel roundness plays a role in rating the size of objects with the example of [y] and [i].)

The statistical evidence shows the difference; we have, however, only one possible explanation for the phenomenon: it may be that the sensitivity of Hungarian participants to vowel harmony could explain the results (cf. Hulst 2019; Törkenczy 2011). As shown, vowel harmony systems, may amplify these size-related symbolic effects by creating uniform vowel qualities within a word (Akinbo 2021). For example, Korean sound-symbolic ideophones exhibit vowel harmony that correlates with iconic size cues, reinforcing the association between front vowels and smallness versus back vowels and largeness (cf. Kwon 2018).

Such patterns suggest that vowel harmony may play a role in shaping and enhancing perceptual size judgments in sound symbolism. Vowel harmony may interact with size rating strategies in sound symbolism by reinforcing coherent perceptual mappings between vowel quality and semantic dimensions such as size. In many languages, front vowels are associated with smallness and lightness, while back vowels evoke largeness and heaviness. When vowel harmony enforces consistent frontness or backness within a word, it can amplify these associations by creating a unified acoustic impression aligned with a particular size category. This effect highlights how phonological systems can influence listeners' symbolic interpretations of sound–meaning relationships.

Akita et al. (2013) and Dingemanse (2012) show that vowel harmony plays a role in the case of ideophones. Tsur (2006) points out that in Hungarian, vowel harmony is present in size-sound symbolic pairs. We know that infants who have no experience in vowel harmony are nevertheless sensitive towards it (Mintz et al. 2018). Thus, we may argue that the better-perceived rounded-unrounded contrast may be the result of the more elaborate and complex learning experience of Hungarians due to their experience in learning how vowel harmony works.

The strategies named by the participants provide us with cues as to why size rating is connected to sound symbolism: there are several different strategies for size ratings of artificial words and these strategies are not exclusive but exist along with each other, meaning that the same participant may use two or three different strategies in parallel in rating the size of objects. Because of the relatively low number of participants who named interesting substrategies (e.g., associations, CDS), it was not possible to statistically analyze each strategy and substrategy, and show whether they actually affected size ratings. That participants consciously named these strategies shows, however, that it is worth asking them to name the strategies they used, and then evaluating the size ratings also in light of the named strategies.

4.3. General discussion

Winter (2025) examines the specific analogies driving sound-shape and sound-size correspondences in spoken language, posing the question of whether these links are driven by how phonemes sound (auditory properties) or how they are produced (articulatory properties)? He argues, based on empirical research, that acoustic analogies alone frequently suffice to explain these phenomena, making articulatory explanations unnecessary. Ekström (2022), however, concludes, reviewing current literature, that current explanations of size-sound symbolism alone may not explain inconsistent findings in empirical research connected to size-sound symbolism. Imai et al. (2025) however, suggest that linguistic symbols may originate from iconic associations across sensory modalities – particularly through the correspondence between oral gestures during speech production and embodied sensory experiences of the world.

Based on the earlier-described causal directionality of Sidhu (2025), we assume that experience may influence the actual effect of size ratings. Kovács, Németh & Elsen (2024) showed that participants' characteristics may modify (intensify or weaken) the effect of phonetic features on size ratings. Thus, size ratings may depend on individual characteristics. McLean, Dunn & Dingemanse (2023) also argue, that sound symbolic assessments may be connected to individual experiences.

Connected to the above results, we argue that previous experiences may also influence the size assessment of artificial words: in our case, earlier exposure (as a child) to CDS may partly modify size ratings.

We see from the above results that in at least one aspect of sound symbolism, namely that of size ratings, several different factors play a role. For the results to be generalizable, we must consider and account for all factors that may influence size rating, and use statistical methods to ensure generalizability (cf. Styles & Gawne 2017). Based on the above considerations and results, the actual size ratings are influenced by multiple factors (Figure 1).

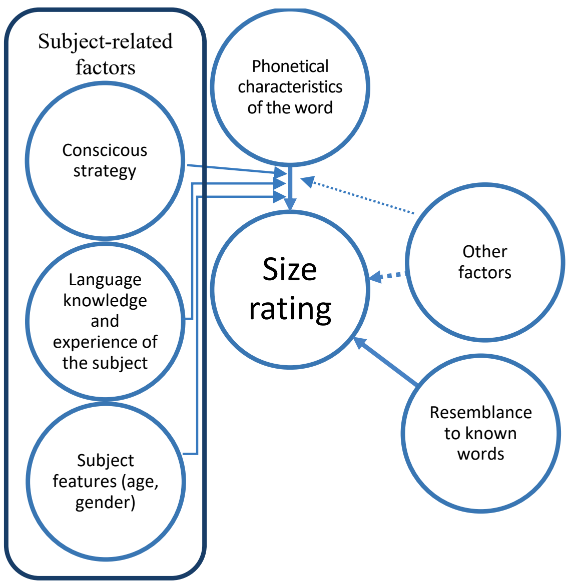

Figure 1. Factors directly or indirectly affecting size rating.

We assume that size rating is a complex phenomenon, in which several factors interact with each other. These factors may be individuum-specific. These complex interactions may explain why some results on size sound symbolism are inconsistent or contradictory. As a result of our findings, we suggest that researchers take into account at least above factors when presenting research results on sound symbolism.

4.4. Limitations

Current study has several limitations.

1. One limitation is that CDS words were categorized based on native-speaker evaluations, as there is no comprehensive list of CDS words with reduplication. Reduplication, a common feature in CDS, aids infants' language learning by simplifying words. Since no standardized inventory of CDS words exist (existing norms contain no, or only a small number of reduplicative words), the classification relies on subjective judgments of the authors, highlighting a limitation of the study. Future work should address this limitation by conducting perceptual norming studies in which independent native speakers rate the CDS-likeness of reduplicative forms. As an alternative, for future research it would be it would be helpful to develop standardized inventories of child-directed speech (CDS) corpora to facilitate cross-linguistic comparisons.
2. A second limitation concerns the treatment of the 1–4 size ratings as a continuous outcome rather than modeling them ordinally; while this choice was made to ensure comparability with prior work, and the main effects are large and consistent across models and samples, making it unlikely that an ordinal analysis would change the substantive conclusions, future studies could nevertheless validate these findings using ordinal or multinomial mixed-effects models.
3. A further limitation of the study is that coding of participant strategies was conducted by a single researcher with verification but no independent recoding, which may introduce bias. Additionally, categories were developed based on self-reports and refined during coding due to the absence of prior research explicitly analyzing size-related strategies in sound symbolism. Future research may address this limitation by using two or more coders or a prearranged set of categories for strategies. In this case, participants only need to indicate whether a given strategy influenced their decision.

5. Conclusion

The current paper sought to answer two less-researched contexts of size ratings: namely, whether child-directed speech may influence results and whether reported size rating strategies are congruent with the actually used strategies.

We defined a feature denoting whether the 2-syllable version of the given stimulus word is an element of CDS in the given country. We found that this feature has a negative effect on the size rating ("diminutive") and that this effect is strong and significant in both participant groups, even after adjusting for other phonetic features. That is, having two words with the same phonetic features, the one resembling a CDS-word tends to be rated as "smaller". However, the inclusion of this factor did not change the strength and significance of the phonetic features.

Comparing conscious (named) strategies with those that were actually used, we showed that at least one named strategy is congruent with the strategies actually employed. Among the several conscious (named) strategies examined, we found that the one based on vowels is congruent with the strategies actually used.

The effect of all factors that we found to be significant was quite strong, and their strength is comparable to the effect of the phonetic features that we found to be relevant in size rating. This in itself demonstrates the linguistic relevance of the issues we raise.

References

Akinbo, Samuel K. 2021. Featural affixation and sound symbolism in Fungwa. *Phonology* 38(4). 537–569. https://doi.org/10.1017/s0952675721000324.

Akinbo, Samuel K. 2023. Reduplication, Repetition and Sound Symbolism in Fungwa. *Phonological Data and Analysis* 5(2). 1–24. https://doi.org/10.3765/pda.v5art2.72.

Akita, Kimi, Mutsumi Imai, Noburo Saji, Katerina Kantartzis & Sotaro Kita. 2013. Mimetic vowel harmony. In Bjarke Frellesvig & Peter Sells (eds.), *Japanese/Korean Linguistics 20*, 115–129. Stanford: CSLI.

Baayen, Harald R. 2008. *Analyzing linguistic data: A practical introduction to statistics using R.* Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9780511801686.

Barr, Dale J., Roger Levy, Christoph Scheepers & Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68(3). 255–278. https://doi.org/10.1016/j.jml.2012.11.001.

Basnak, Johanna & Ota, Mitsuhiko. 2024. Learnability Advantage of Segmental Repetitions in Word Learning. *Language and Speech* 67(4). 1093–1120. https://doi.org/10.1177/00238309231223909.

Bates, Douglas, Reinhold Kliegl, Shravan Vasishth & Harald Baayen. 2015. *Parsimonious Mixed Models*. arXiv:1506.04967. https://doi.org/10.48550/arXiv.1506.04967.

Benczes, Réka 2018. *Rhyme over Reason*. Cambridge: Cambridge University Press.

Blasi, Damián E., Søren Wichmann, Harald Hammarström, Peter Stadler & Morten H. Christiansen. 2016. Sound–Meaning association biases evidenced across thousands of languages. *Proceedings of the National Academy of Sciences* 113(39). 10818–10823. https://doi.org/10.1073/pnas.1605782113.

Burnham, Denis, Christine Kitamura & Ute Vollmer-Conna. 2002. What's new, pussycat? On talking to babies and animals. *Science* 296(5572). 1435. https://doi.org/10.1126/science.1069587.

Campisi, Emanuela & Asli Özyürek. 2013. Iconicity as a communicative strategy: Recipient design in multimodal demonstrations for adults and children. *Journal of Pragmatics* 47(1). 14–27. https://doi.org/10.1016/j.pragma.2012.12.007.

Cathcart, Chundra A. 2024. Multiple evolutionary pressures shape identical consonant avoidance in the world's languages. *Proceedings of the National Academy of Sciences* 121(27). e2316677121. https://doi.org/10.1073/pnas.2316677121.

Cooper, Robin Panneton & Richard N. Aslin. 1990. Preference for infant-directed speech in the first month after birth. *Child Development* 61 (5). 1584–1595. https://doi.org/10.2307/1130766.

Ćwiek, Aleksandra, Rémi Anselme, Dan Dediu, Susanne Fuchs, Shigeto Kawahara, Grace E. Oh, Jing Paul, Marcus Perlman, Caterina Petrone, Sabine Reiter, Rachid Ridouane, Jochen Zeller & Bodo Winter. 2024. The Alveolar Trill Is Perceived as Jagged/Rough by Speakers of Different Languages. *The Journal of the Acoustical Society of America* 156(5). 3468–3479. https://doi.org/10.31234/osf.io/cv95m.

Ćwiek, Aleksandra & Susanne Fuchs. 2025. The Scope and Limits of Iconic Prosody: Head Angle Predicts f0 Changes While Object Size Effects Are Absent. *Open Mind* 9. 1959–1981. https://doi.org/10.1162/OPMI.a.251.

D'Anselmo, Anita, Giulia Prete, Przemysław Zdybek, Luca Tommasi & Alfredo Brancucci. 2019. Guessing meaning from word sounds of unfamiliar languages: A cross-cultural sound symbolism study. *Frontiers in Psychology* 10. 593. https://doi.org/10.3389/fpsyg.2019.00593.

Diffloth, Gérard. 1994. I: Big, a: Small. In Leanne Hinton, Joanna Nichols & John J. Ohala (eds.), *Sound Symbolism,* 107–114. Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9780511751806.008.

Dingemanse, Mark. 2012. Advances in the cross-linguistic study of ideophones. *Language and Linguistics Compass* 6(10). 654–672. https://doi.org/10.1002/lnc3.361.

Dressler, Wolfgang U., Laura E. Lettner & Katharina Korecky-Kröll. 2012. Acquisition of German diminutive formation and compounding in a comparative perspective. In Kiefer Ferenc, Ladányi Mária & Siptár Péter (eds.), *Current Issues in Morphological Theory: (Ir)regularity, analogy and frequency. Selected papers from the 14th International Morphology Meeting, Budapest, 13–16 May 2010.*, 237–264. Amsterdam–Philadelphia: John Benjamins. https://doi.org/10.1075/cilt.322.11dre.

Dupuis, Marc, Emanuele Meier & Félix Cuneo. 2019. Detecting computer-generated random responding in questionnaire-based data: A comparison of seven indices. *Behavior Research Methods* 51(5). 2228–2237. https://doi.org/10.3758/s13428-018-1103-y.

Ekström, Axel G. 2022. What's next for size-sound symbolism? *Frontiers in Language Sciences* 1. 1046637. https://doi.org/10.3389/flang.2022.1046637.

Elsen, Hilke. 2008. *Phantastische Namen. Die Namen in Science Fiction und Fantasy zwischen Arbitrarität und Wortbildung*. Tübingen: Narr. https://doi.org/10.24053/9783823373964.

Elsen, Hilke. 2016. *Einführung in die Lautsymbolik*. Berlin: Erich Schmidt. https://doi.org/10.37307/b.978-3-503-16790-6.

Elsen, Hilke. 2017. The two meanings of sound symbolism. *Open Linguistics* 3(1). 489–497. https://doi.org/10.1515/opli-2017-0024.

Elsen, Hilke. 2018. Some names are more equal than others – The sound symbolic value of new names. *Cahiers de lexicologie et normes* 113(2). 79–94.

Elsen, Hilke. 2019. Lautsymbolik in phantastischen Sprachen. *Wirkendes Wort* 69(1). 103–119.

Elsen, Hilke, Renáta Németh & László Kovács. 2021. The sound of size revisited – New insights from a German–Hungarian comparative study on sound symbolism. *Language Sciences* 85. 101360. https://doi.org/10.1016/j.langsci.2021.101360.

Erben Johansson, Niklas, Andrey Anikin & Nikolay Aseyev. 2019. Color sound symbolism in natural languages. *Language and Cognition* 12(1). 56–83. https://doi.org/10.1017/langcog.2019.35.

Erben Johansson, Niklas, Andrey Anikin, Gerd Carling & Arthur Holmer. 2020. The typology of sound symbolism: Defining macro-concepts via their semantic and phonetic features. *Linguistic Typology* 24(2). 253–310. https://doi.org/10.1515/lingty-2020-2034.

Ferber, Michael. 2019. Onomatopoeia and Sound Symbolism. In Michael Ferber (ed.), *Poetry and Language: The Linguistics of Verse,* 86–103. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108554152.004.

Fónagy, Iván. 1961. Communication in poetry. *Word* 17. 194–218. https://doi.org/10.1080/00437956.1961.11659754.

Foulkes, Paul, Gerard J. Docherty & Dominic Watt. 2005. Phonological Variation in Child-Directed Speech. *Language* 81(1). 177–206. https://doi.org/10.1353/lan.2005.0018.

Garmann, Nina Gram, Pernille Hansen, Hanne Gram Simonsen & Kristian Emil Kristoffersen. 2019. The Phonology of Children's Early Words: Trends, Individual Variation, and Parents' Accommodation in Child-Directed Speech. *Frontiers in Communication* 4. 10. https://doi.org/10.3389/fcomm.2019.00010.

Golinkoff, Roberta Michnick, Dilara Deniz Can, Melanie Soderstrom & Kathy Hirsh-Pasek. 2015. (Baby)talk to me: The social context of infant-directed speech and its effects on early language acquisition. *Current Directions in Psychological Science* 24(5). 339–344. https://doi.org/10.1177/0963721415595345.

Han, Mengru, Yiqi Nie & Yan Gu. 2024. Bridging Word and World: Vocal Iconicity in Chinese Child-Directed Speech and Child Production. *Proceedings of the Annual Meeting of the Cognitive Science Society 46*. https://escholarship.org/uc/item/3mz2j3z0.

Heinsoo, Heinike & Eva Saar. 2015. Sound symbolism of expressive verbs in Finnic languages (Estonian, Finnish, Ingrian, Votic). *Eesti Ja Soome-Ugri Keeleteaduse Ajakiri. Journal of Estonian and Finno-Ugric Linguistics* 6(1). 55–74. https://doi.org/10.12697/jeful.2015.6.1.04.

Hinton, Leanne, Johanna Nichols & John Jerome Ohala. (eds.) 1994. *Sound symbolism*. Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9780511751806.

Hulst, Harry van der. 2019. Vowel Harmony in Hungarian: A Comparison of Segmental and Autosegmental Analyses. In Harry van der Hulst & Norval Smith (eds.), *Advances in Nonlinear Phonology*, 267–304. Berlin, Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110869194-015.

Imai, Mutsumi & Sotaro Kita. 2014. The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences* 369(1651). 20130298. https://doi.org/10.1098/rstb.2013.0298.

Imai, Mutsumi, Sotaro Kita, Miho Nagumo & Hiroyuki Okada. 2008. Sound symbolism facilitates early verb learning. *Cognition* 109(1). 54–65. https://doi.org/10.1016/j.cognition.2008.07.015.

Imai, Mutsumi, Sotaro Kita, Kimi Akita, Noburo Saji, Masato Ohba & Miki Namatame. 2025. Does sound symbolism need sound?: The role of articulatory movement in detecting iconicity between sound and meaning. *The Journal of the Acoustical Society of America* 157(1). 137–148. https://doi.org/10.1121/10.0034832.

Joo, Ian. 2020. Phonosemantic biases found in Leipzig-Jakarta lists of 66 languages. *Linguistic Typology* 24(1). 1–12. https://doi.org/10.1515/lingty-2019-0030.

Jun, Yesel & Hyunju Lee. 2022. A sound brand identity design: The interplay between sound symbolism and typography on brand attitude and memory. *Journal of Retailing and Consumer Services* 64. 102724. https://doi.org/10.1016/j.jretconser.2021.102724.

Kantartzis, Katerina. 2011. *Children and adults' understanding and use of sound-symbolism in novel words*. University of Birmingham. PhD thesis.

Kawahara, Shigeto. 2020. A wug-shaped curve in sound symbolism: the case of Japanese Pokémon names. *Phonology* 37(3). 383–418. https://doi.org/10.1017/s0952675720000202.

Kawahara, Shigeto & Canaan Breiss. 2021. Exploring the nature of cumulativity in sound symbolism: Experimental studies of Pokémonastics with English speakers. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 12(1). 3. https://doi.org/10.5334/labphon.

Kawahara, Shigeto & Gakuji Kumagai. 2019. Expressing evolution in Pokémon names: Experimental explorations. *Journal of Japanese Linguistics* 35(1). 3–38. https://doi.org/10.1515/jjl-2019-2002.

Kawahara, Shigeto, Mahayana C. Godoy & Gakuji Kumagai. 2020. Do Sibilants Fly? Evidence from a Sound Symbolic Pattern in Pokémon Names. *Open Linguistics* 6(1). 386–400. https://doi.org/10.1515/opli-2020-0027.

Kawahara, Shigeto, Aiko Noto & Gakuji Kumagai. 2018. Sound Symbolic Patterns in Pokémon Names. *Phonetica* 75(3). 219–244. https://doi.org/10.1159/000484938.

Kempe, Vera, Patricia J. Brooks & Steven Gillis. 2005. Diminutives in child-directed speech supplement metric with distributional word segmentation cues. *Psychonomic Bulletin & Review* 12(1). 145–151. https://doi.org/10.3758/bf03196360.

Klink, Richard R. & Lan Wu. 2014. The role of position, type, and combination of sound symbolism imbeds in brand names. *Marketing Letters* 25(1). 13–24. https://doi.org/10.1007/s11002-013-9236-3.

Knoeferle, Klemens, Jixing Li, Emanuela Maggioni & Charles Spence. 2017. What drives sound symbolism? Different acoustic cues underlie sound-size and sound-shape mappings. *Scientific Reports* 7. 5562. https://doi.org/10.1038/s41598-017-05965-y.

Köhler, Wolfgang. 1929. *Gestalt Psychology. An Introduction to New Concepts in Modern Psychology*. Liveright, New York/Toronto.

Kouwenberg, Silvia & Darlene LaCharité. 2005. Less is more: Evidence from diminutive reduplication in Caribbean Creole languages. In Bernhard Hurch (ed.), *Studies on Reduplication*, 533–546. Berlin & New York: Mouton de Gruyter. https://doi.org/10.1515/9783110911466.533.

Kovács, László, Renáta Németh & Hilke Elsen. 2024. Often Overlooked Aspects of Sound Symbolism: The Influence of Participants' Characteristics on Size Ratings. *Language and Speech* 67(3). 772–792. https://doi.org/10.1177/00238309231193631.

Kwon, Nahyun. 2018. Iconicity correlated with vowel harmony in Korean ideophones. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 9(1). 1. https://doi.org/10.5334/labphon.53.

Laalo, Klaus & Reili Argus. 2020. Linguistic recycling in language acquisition. *Linguistic Recycling* 33. 86–103. https://doi.org/10.1075/aila.00031.laa.

Laing, Catherine. 2019. A role for onomatopoeia in early language: Evidence from phonological development. *Language and Cognition* 11(2). 173–187. https://doi.org/10.1017/langcog.2018.23.

Laing, Catherine, Ghada Khattab, Shayne Sloggett & Tamar Keren-Portnoy. 2025. Size sound symbolism in mothers' speech to their infants. *Journal of Child Language* 52(4). 739–761. https://doi.org/10.1017/S0305000921000799.

Lockwood, Gwilym & Mark Dingemanse. 2015. Iconicity in the lab: a review of behavioral, developmental, and neuroimaging research into sound-symbolism. *Frontiers in Psychology* 6. https://doi.org/10.3389/fpsyg.2015.01246.

McLean, Bonnie, Michael Dunn & Mark Dingemanse. 2023. Two measures are better than one: combining iconicity ratings and guessing experiments for a more nuanced picture of iconicity in the lexicon. *Language and Cognition* 15(4). 716–739. http://doi.org/10.1017/langcog.2023.9.

Mintz, Toben H., Rachel L. Walker, Ashlee Welday & Celeste Kidd. 2018. Infants' sensitivity to vowel harmony and its role in segmenting speech. *Cognition* 171. 95–107. https://doi.org/10.1016/j.cognition.2017.10.020.

Motamedi, Yasamin, Hannah Little, Alan Nielsen & Justin Sulik 2019. The iconicity toolbox: Empirical approaches to measuring iconicity. *Language and Cognition* 11(2). 188–207. https://doi.org/10.1017/langcog.2019.14.

Motoki, Kosuke, Sayo Iseki & Abhishek Pathak 2025. Sounds cute: Exploring the role of sound reduplication in brand names. *International Journal of Research in Marketing* 42(4) Part B. 1305–1322. https://doi.org/10.1016/j.ijresmar.2025.02.002.

Nielsen, Alan K. S. & Mark Dingemanse. 2021. Iconicity in Word Learning and Beyond: A Critical Review. *Language and Speech* 64(1). 52–72. https://doi.org/10.1177/0023830920914339.

Nuckolls, Janis B. 1996. *Sounds Like Life: Sound-Symbolic Grammar, Performance, and Cognition in Pastaza Quechua*. New York, Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780195089851.001.0001.

Ota, Mitsuhiko & Barbara Skarabela. 2016. Reduplicated words are easier to learn. *Language Learning and Development* 12(4). 380–397. https://doi.org/10.1080/15475441.2016.1165100.

Ota, Mitsuhiko & Barbara Skarabela. 2018. Reduplication facilitates early word segmentation. *Journal of Child Language* 45(1). 204–218. https://doi.org/10.1017/S0305000916000660.

Outters, Vivien, Melanie S. Schreiner, Tanya Behne & Nivedita Mani. 2020. Maternal input and infants' response to infant-directed speech. *Infancy* 25(4). 478–499. https://doi.org/10.1111/infa.12334.

Pejovic, Jovana & Monika Molnar. 2017. The development of spontaneous sound-shape matching in monolingual and bilingual infants during the first year. *Developmental Psychology* 53(3). 581. https://doi.org/10.1037/dev0000237.

Peña, Marcela, Jacques Mehler & Marina Nespor. 2011. The role of audiovisual processing in early conceptual development. *Psychological Science* 22(11). 1419–1421. https://doi.org/10.1177/0956797611421791.

Perry, Lynn K., Marcus Perlman, Bodo Winter, Dominic W. Massaro & Gary Lupyan. 2018. Iconicity in the speech of children and adults. *Developmental Science* 21(3). e12572. https://doi.org/10.1111/desc.12572.

Perry, Lynn K., Stephanie A. Custode, Regina M. Fasano, Brittney M. Gonzalez & Jordyn D. Savy. 2021. What Is the Buzz About Iconicity? How Iconicity in Caregiver Speech Supports Children's Word Learning. *Cognitive Science* 45(4). e12976. https://doi.org/10.1111/cogs.12976.

Pisanski, Katarzyna, Sari G. E. Isenstein, Kelyn J. Montano, Jillian J. M. O'Connor & David R. Feinberg. 2017. Low is large: Spatial location and pitch interact in voice-based body size estimation. *Attention, Perception, & Psychophysics* 79(4). 1239–1251. https://doi.org/10.3758/s13414-016-1273-6.

Rabe-Hesketh, Sophia & Andres Skrondal. 2012. *Multilevel and Longitudinal Modeling Using Stata.* College Station: Stata Press.

Ruiter, Laura E. de, Anna L. Theakston, Silke Brandt & Elena V.M. Lieven. 2018. Iconicity affects children's comprehension of complex sentences: The role of semantics, clause order, input and individual differences. *Cognition* 171. 202–224. https://doi.org/10.1016/j.cognition.2017.10.015.

Saji, Noburo, Kimi Akita, Katerina Kantartzis, Sotaro Kita & Mutsumi Imai. 2019. Cross-linguistically shared and language specific sound symbolism in novel words elicited by locomotion videos in Japanese and English. *PLOS ONE* 14(7). e0218707. https://doi.org/10.1371/journal.pone.0218707.

Saxton, Matthew. 2009. The Inevitability of Child Directed Speech. In Susan H. Foster-Cohen (ed.), *Language Acquisition,* 62–86. London: Palgrave Macmillan. https://doi.org/10.1057/9780230240780_4.

Shinohara, Kazuko & Shigeto Kawahara. 2010. A cross-linguistic study of sound symbolism: The image of size. *Berkeley Linguistics Society* 36. 396–410. https://doi.org/10.3765/bls.v36i1.3926.

Shinohara, Kazuko, Shigeto Kawahara & Hideyuki Tanaka. 2020. Visual and Proprioceptive Perceptions Evoke Motion-Sound Symbolism: Different Acceleration Profiles Are Associated With Different Types of Consonants. *Frontiers in Psychology* 11. https://doi.org/10.3389/fpsyg.2020.589797.

Shrum, L. J., Tina M. Lowrey, Deborah Luna, Dawn Lerman & Min Liu. 2012. Sound symbolism effects across languages: Implications for global brand names. *International Journal of Research in Marketing* 29(3). 275–279. https://doi.org/10.1016/j.ijresmar.2012.03.002.

Sidhu, David M. 2025. Sound Symbolism in the Lexicon: A Review of Iconic-Systematicity. *Language and Linguistics Compass* 19(1). 1–9. https://doi.org/10.1111/lnc3.70006.

Sidhu, David M. & Penny M. Pexman. 2018. Five mechanisms of sound symbolic association. *Psychonomic Bulletin & Review* 25. 1619–1643. https://doi.org/10.3758/s13423-017-1361-1.

Sidhu, David M., Gabriella Vigliocco & Penny M. Pexman. 2019. Effects of iconicity in lexical decision. *Language and Cognition* 12(1). 164–181. https://doi.org/10.1017/langcog.2019.36.

Snijders, Tom A. B. & Roel J. Bosker. 2012. *Multilevel analysis: an introduction to basic and advanced multilevel modeling.* London: Sage.

Styles, Suzy J., & Lauren Gawne. 2017. When Does Maluma/Takete Fail? Two Key Failures and a Meta-Analysis Suggest That Phonology and Phonotactics Matter. *I-Perception* 8(4). 2041669517724807. https://doi.org/10.1177/2041669517724807.

Thompson, Bill, Marcus Perlman, Gary Lupyan, Zed Sevcikova Sehyr & Karen Emmorey. 2020. A data-driven approach to the semantics of iconicity in American Sign Language and English. *Language and Cognition* 12(1). 182–202. https://doi.org/10.1017/langcog.2019.52.

Törkenczy, Miklós. 2011. Hungarian Vowel Harmony. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology,* 2963–2990. Oxford: Wiley-Blackwell. https://doi.org/10.1002/9781444335262.wbctp0123.

Tsur, Reuven. 1992. *What Makes Sound Patterns Expressive: The Poetic Mode of Speech-Perception (in English)*. Durham: Duke University Press. https://doi.org/10.1215/9780822378365.

Tsur, Reuven. 2006. Size–sound symbolism revisited. *Journal of Pragmatics* 38(6). 905–924. https://doi.org/10.1016/j.pragma.2005.12.002.

Tsur, Reuven. 2012. Playing *by Ear and the Tip of the Tongue.* Amsterdam/Philadelphia: John Benjamins.

Usnadze, Dimitri. 1924. Ein experimenteller Beitrag zum Problem der psychologischen Grundlagen der Namengebung. *Psychological Research/Psychologische Forschung* 5(1). 24–43.

Vainio, Lari. 2021. Magnitude sound symbolism influences vowel production. *Journal of Memory and Language* 118. 104213. https://doi.org/10.1016/j.jml.2020.104213.

Winter, Bodo. 2025. The size and shape of sound: The role of articulation and acoustics in iconicity and crossmodal correspondences. *The Journal of the Acoustical Society of America* 157(4). 2636–2656. https://doi.org/10.1121/10.0036362.

Winter, Bodo & Marcus Perlman. 2021. Size sound symbolism in the English lexicon. *Glossa* 6(1). 79. https://doi.org/10.5334/gjgl.1646.

Wissemann, Heinz. 1954. *Untersuchungen zur onomatopoiie, I. Teil: Die sprach-psychologischen Versuche.* Heidelberg: C. Winter.

Wu, Lan, Richard R. Klink & Jiansheng Guo. 2013. Creating gender brand personality with brand names: The effects of phonetic symbolism. *Journal of Marketing Theory and Practice* 21(3). 319–329. https://doi.org/10.2753/mtp1069-6679210306.

Kokkuvõte. László Kovács, Renáta Németh, Hilke Elsen: Helisümbolism suuruse tajumisel: lapsele suunatud kõne ja strateegia efektide koosmõju. Helisümbolism – foneetiliste omaduste ja tähenduse süstemaatiline seos – mängib olulist rolli suuruse hindamisel, kus teatud häälikud kujundavad suuruse tajumist. Käesolevas artiklis uuritakse kahte varem vähe uuritud suuruse ja helisümbolismi mõõdet: (1) lastele suunatud kõne mõju suuruse hindamisele ja (2) osalejate enda antud hinnangute ja tegelike hindamisstrateegiate vastavust. Esimeses osas antakse lühike ülevaade helisümbolismi teoreetilisest taustast ja selle seosest suuruse tajumisega. Teises osas esitatakse olemasoleva andmestiku uus analüüs, tuginedes ungari ja saksa keele kõnelejate empiirilistele andmetele. Meie tulemused näitavad tugevat pisendavat mõju: lastele suunatud kõne vormidega sarnanevaid tehissõnu hinnatakse järjekindlalt väiksemaks, isegi pärast teiste foneetiliste muutujate arvessevõtmist. Lisaks leidsime, et vähemalt üks selgesõnaliselt nimetatud strateegia – vokaalidel põhinevad hinnangud – peegeldab täpselt strateegiaid, mida osalejad tegelikult kasutavad.

Märksõnad: heli sümbolism, suuruse heli sümbolism, suuruse hindamine, lapsele suunatud kõne

Appendix 1.

Table 6. The 56 CVCVCV stimuli in an alphabetical order. Participants viewed the stimuli in one of ten randomized orders and provided ratings for each stimulus.

bababa	*fafafa*	*kakaka*	*Tatata*
bebebe	*fefefe*	*kekeke*	*Tetete*
bibibi	*fififi*	*kikiki*	*tititi*
bobobo	*fofofo*	*kokoko*	*tototo*
böböbö	*föföfö*	*kökökö*	*tötötö*
bububu	*fufufu*	*kukuku*	*tututu*
bübübü	*füfüfü*	*kükükü*	*tütütü*
dadada	*gagaga*	*papapa*	*vavava*
dedede	*gegege*	*pepepe*	*veveve*
dididi	*gigigi*	*pipipi*	*vivivi*
dododo	*gogogo*	*popopo*	*vovovo*
dödödö	*gögögö*	*pöpöpö*	*vövövö*
dududu	*gugugu*	*pupupu*	*vuvuvu*
düdüdü	*gügügü*	*püpüpü*	*vüvüvü*

Appendix 2.

Table 7. Identified strategies for size ratings, based on self-reported strategies.

1. Vowel	11. Vowel light / dark 12. Vowel length 13. Umlaut
2. Consonant	21. Voiced / Voiceless Consonants
3. Similar sounding words	31. Initial letter 32. Child language words 33. Belittlement/minimalization 34. Similar meaning
4. Special strategies	401. Synesthesia 402. Tune to the word 403. Word length 404. Imagined speaker 405. Associations 406. Mouth opening 407. Position in vocal tract 408. Noises 409. Volume 410. Imaginary gestures, facial expressions 411. Interplay of consonants and vowels 412. Stress 413. Pitch
5. General information to sounds	
6. Nothing specific	
7. Mixed strategies (may include a combination of above; not clear based on named strategy)	71. Sound (spoken) 72. Letter (written) 73. Lx (some known language) 74. L1 (mother tongue) 75. L1 + Lx (mother tongue and other languages)

Appendix 3.

Table 8. Correlation between the variable CDS and the phonetic features, each of which is binary 0/1. Height is encoded in two dummies: middle and high. Significance at the 0.001 level is denoted by '***'.

	backness	roundness	voicing	articulation	height= high	height= middle
Hungarian sample	0.05***	–0.1643***	0	0.2***	0.05***	–0.0913***
German sample	0.235***	–0.235***	–0.0626***	0.1808***	–0.1446***	–0.1446***