DE GRUYTER Open Linguistics 2023; 9: 20220257

Research Article

Dominic Schmitz*, Defne Cicek, Anh Kim Nguyen, and Daniel Rottleb

Cuteness modulates size sound symbolism at its extremes

https://doi.org/10.1515/opli-2022-0257
received May 19, 2023; accepted November 13, 2023

Abstract: Despite the rapidly growing body of research on sound symbolism, one issue that remains understudied is whether different types of sensory information interact in their sound symbolic effects. The experimental study reported here consisted of two tasks and focused on one such potential interaction: size associations and cuteness. First, a forced-choice task was conducted in which size ratings were elicited for pseudowords containing different vowels and consonants. The pseudowords were introduced as names of alien creatures, which were used as visual stimuli in the experiment. Second, the cuteness of alien creatures was assessed in a judgement task. Both tasks were completed by the same group of German speakers. In line with previous research, /aː/ was associated with largeness and /iː/ was associated with smallness. Further, we found that cuteness modulates size associations in /aː/ and /iː/. For /aː/ judged size increased, while for /iː/ judged size decreased with increasing cuteness. Regarding consonants, we found that /ʁ/ evoked higher size associations than other consonants under investigation. Interactions of cuteness and consonants did not reach significance. Our findings call for the integration of other possible factors and features that might show sound symbolic effects or interactions with such in sound symbolism research.

Keywords: sound symbolism, sound-size correspondence, cuteness, infant schema, German

1 Introduction

Sound symbolism is a specific form of cross-modal correspondence: Certain sounds are associated with certain sensory information. Such sounds, then, are related to meaning, contra the widely acknowledged Saussurean arbitrariness of speech sounds (De Saussure 1916). Sound symbolic patterns and influences have, for instance, been found in consonants and vowels associated with shape (Ahlner and Zlatev 2010, Bremner et al. 2013, Ćwiek et al. 2022; D'Onofrio 2013, Kawahara and Shinohara 2012, Köhler 1929, Maurer et al. 2006, Nielsen and Rendall 2013, Ramachandran and Hubbard 2001, Westbury et al. 2018) and in naming patterns (Berlin 2006, Kawahara 2012, Köhler 1929, Perfors 2004, Ramachandran and Hubbard 2001, Sapir 1929). One of the most frequently researched types of sound symbolism is size sound symbolism, i.e. the association of speech sounds and size. Indeed, Sapir (1929) already uncovered that speakers of English and Chinese apparently associate the open vowel /a/ with largeness, while they associate the close vowel /i/ with smallness. More recent work on size sound symbolism across languages largely is in line with Sapir's findings (Berlin 1995, 2006, Blasi et al. 2016, Johansson 2017, Erben Johansson et al. 2020, Newman 1933, Shinohara and Kawahara 2010, Thompson and Estes 2011, Westbury et al. 2018; see Diffloth 1994 on the directionality of the effect). For consonants, it has been

*** Corresponding author: Dominic Schmitz,** Department of English Language and Linguistics, Heinrich-Heine-Universität Düsseldorf, 40225 Düsseldorf, Germany, e-mail: dominic.schmitz@uni-duesseldorf.de
Defne Cicek, Daniel Rottleb: Department of English Language and Linguistics, Heinrich-Heine-Universität Düsseldorf, 40225 Düsseldorf, Germany
Anh Kim Nguyen: Department of Linguistics, Heinrich-Heine-Universität Düsseldorf, 40225 Düsseldorf, Germany

found that largeness is associated with voiced obstruents (Newman 1933, Hamano 1986, Shinohara and Kawahara 2010, Kawahara and Shinohara 2012).

Even though research on sound symbolism is concerned with a variety of different cross-modal correspondences, thus far, potential interactions between different types of sensory information have been neglected. Take, for example, the popular pair of pseudowords coined by Ramachandran and Hubbard (2001), following the seminal work by Köhler (1929): *bouba* and *kiki*. Commonly, these pseudowords are used in experiments in combination with a round shape (*bouba*) and a spiky shape (*kiki*) to demonstrate shape sound symbolism: *bouba* is more often associated with a round shape, and *kiki* is more often associated with a spiky shape by participants (e.g. Ćwiek et al. 2022). In respective experiments, both shapes are shown in roughly the same size. But what if both shapes would differ in size? Let us imagine an experiment with four visual stimuli: a small spiky shape, a large spiky shape, a small round shape, and a large round shape. Each participant of this imaginary experiment would be presented with a small and a large shape, each differing in their shape quality. This leaves us with two conditions: Participants see either a small round shape and a big spiky shape or a big round shape and a small spiky shape. The participants' task is to choose which of the two shown shapes fit *bouba* (or *kiki*) best. Taking into account only findings on shape sound symbolism, we would assume that in both conditions the round shapes will be more often associated with *bouba* than with *kiki*. If we were to take into account only findings on size sound symbolism, however, we would assume that the smaller-sized shapes regardless of their shape quality are associated more often with *kiki*, as the word contains two /i/ sounds. On the other hand, the bigger-sized shapes would be associated more often with *bouba*, as the word's /a/ sound is associated with largeness. The question that naturally suggests itself at this point is: Do the two types of sensory information, shape and size, interact or override each other in regard to their sound symbolic effects, and if so, in what way?

A similar question is the focus of the present article. While we will retain size as one sensory information of interest, we will pair it with cuteness instead of shape. Thus far, only little sound symbolism research on cuteness is available. For Japanese, it was found that the bilabials /p/, /b/, /m/, /ɸ/, and /w/ are associated with cuteness (Kumagai and Kawahara 2017, Kumagai 2019, 2020), while for English, only the unaspirated, low-frequency bilabials /b/ and /m/ seem to be associated with cuteness (Kumagai 2020).

Somewhat related to cuteness are findings on other positive traits. Klink (2000) found that positive traits such as friendliness and prettiness are associated with front vowels in American English. Kilpatrick et al. (2023) showed that cross-linguistically so-called friendship values in Pokémon are typically high in those pocket monsters with bilabial plosives in their names. Pokémon with /m/, /d/, and /g/ in their names, on the other hand, come with lower friendship values, while /w/ is associated with neither low nor high friendship values.

What remains is the question of why certain sounds evoke associations of cuteness. Cuteness, as comprised from a biological perspective in the so-called 'infant schema' (e.g. Lehmann et al., 2013; Lorenz 1943), is a fundamental feature of human perception and correlates most prominently with size. During the last decade, the idea of the infant schema and with it the concept of cuteness have been extended beyond visual features to include olfactory and, most importantly, auditory features (Kringelbach et al. 2016). Babbling infants mostly produce high-frequency, vowel-like sounds and labial consonants (MacNeilage and Davis 2000, Kringelbach et al. 2016), and the prototypical idea of an infant can reasonably be related to sucking milk. The latter idea is also brought forward by Kumagai (2020) as a potential explanation for their findings on Japanese bilabials. However, for their more restricted findings on English, this explanation is not sufficient. Thus, they put forward semantic differences between English *cute* and Japanese *kawaii* as a potential reason for their differing findings. While Japanese *kawaii* is restricted to one meaning, i.e. *adorable*, *lovable*, English *cute* may also be understood as *clever* or *cunning*. The latter, in turn, may relate to findings on villainess, which is associated with low-frequency consonants such as /b/ and /m/ (Uno et al. 2020). This, in turn, might strengthen the English associations between cuteness and /b/ and /m/. Further, Kumagai (2020) argues that the aspiration following /p/ might counteract cuteness associations for English, as aspiration is potentially related to strength.

Seeing as cuteness and size are both associated with different consonants and vowels, the present article aims to answer the following research question: Do size and cuteness associations interact when a referent has both traits? The remainder of this article is structured as follows. Section 2 will introduce the methodology of the present investigation, including the choice of stimuli, the procedure of the experiment used to elicit the required data, and the statistical analysis. The results of the experiment will be presented in Section 3. Section 4 will then conclude this article with a discussion.

2 Methods

To analyse a potential interaction of size and cuteness in sound symbolism, data on these two dimensions were collected in an experiment that consisted of two parts, a forced-choice task and a judgement task. Data on size were elicited in the forced-choice task; data on cuteness were elicited in the judgement task. The experiment was created in OpenSesame (Mathôt et al. 2012) and distributed via the MindProbe server (https://mindprobe.eu/). The stimuli used in the experiment as well as the procedure during the experiment are explained in the following sections. Stimuli, data, and scripts are available at the OSF: https://osf.io/qck9r/.

2.1 Stimuli

2.1.1 Auditory stimuli

Auditory stimuli were used in the forced-choice task for size rating elicitation. The auditory stimuli consisted of recordings of pseudowords to rule out potential effects that real words are known to introduce to experiments: storage effects (e.g. Caselli et al. 2016), frequency effects (e.g. Gahl 2008, Lohmann 2018), and effects of lexical relatedness (e.g. Schriefers et al. 1998). One factor that may affect auditory stimulus processing, even when using pseudowords, is phonological neighbourhood density (PND, Vitevitch and Luce 1998). This potential influence will be controlled for in the analysis by integrating PND as a covariate. Additionally, real words are connected either to real-world perception of size for words that represent non-abstract concepts or to abstract concepts for which size and cuteness as dimensions are either not straightforward to assess or meaningless.

All pseudowords designed for the present study consisted of two syllables, with stress on the first syllable. Both syllables were open, i.e. without coda consonants. Onsets in both syllables were simple and consisted of either /d, f, j, k/, or ʁ/.[1] /d, f, j/, and /k/ were each allowed to constitute the onset of either one or both syllables in a pseudoword. /ʁ/, on the other hand, was only allowed to be the onset of one of the two syllables. This difference between /d, f, j, k/ and /ʁ/ was made because of existing German disyllabic words and their syllable structure. That is, for the earlier four consonants, words containing the same consonant in both onsets and long vowels as nuclei of both open syllables are attested (e.g. *Dodo* /doːdoː/ 'dodo', *FIFA* /fiːfaː/ 'FIFA (acronym)', *Jo-Jo* /joːjoː/ 'yoyo', *Kaki* /kaːkiː/ 'persimon').[2] For /ʁ/, however, there is only one attested word form fitting the disyllabic open syllable structure with simplex onsets, but this word form does not contain two long vowels as nuclei: *rare* /ʁaːʁə/ 'rare (feminine)'. Also, this latter word form is an adjective, whereas the former four words as well as the pseudowords used in the experiment are nouns. As nuclei, the eight long vowels of German /aː, ɛː, eː, iː, oː, øː, uː/ and /yː/ were chosen, and in each pseudoword, both nuclei consisted of the same vowel.[3] Table 1 provides an overview of all potential pseudowords; a list of all 96 spelled-out pseudowords can be found at the OSF.

---

1 One might wonder why, especially in light of previous findings on consonants and cuteness as reported in the Introduction, this set of consonants, which does not contain any bilabial sounds, was chosen. This is due to the circumstance that the present experiment was originally designed to only investigate vowels. We were made aware of the findings on consonants only later by an anonymous reviewer and gratefully integrated them into this article.

2 Note that this later on also led to the exclusion of the data of the two pseudowords /doːdoː/ and /joːjoː/, as they are homophonous to the existing words *Dodo* 'dodo' and *Jo-Jo* 'yoyo.'

3 Note that /ɛː/ and /eː/ are experiencing a merger in German (Kleiner 2011). Depending on regional variation, /eː/ may be produced instead of /ɛː/, calling into question their separate phonemic status. We included both vowels for two reasons. First, the speaker who

Table 1: Phonological make-up of pseudowords

1st syllable			2nd syllable		
C1	V		C2	V	
d	aː	ɛː	d	aː	ɛː
f			f		
j	eː	iː	j	eː	iː
k	oː	øː	k	oː	øː
d, f, j, k	uː	yː	ʁ	uː	yː
ʁ			d, f, j, k		

Consonants (C1, C2) in the same row were allowed to be part of the same pseudoword. Vowels (V) were allowed with all consonants.

During the experiment, pseudoword stimuli were presented as audio cues. For the creation of audio stimuli, a reading list featuring the pseudowords was created. On these lists, items were embedded within the German sentence *Ich habe pseudoword gesagt* 'I said pseudoword'. A trained native speaker of German read the entire reading list aloud for practice before recording the list three times in quiet surroundings. The recordings were sampled at 44.1 kHz, 16 bit.

For each pseudoword, the best of the three recordings was chosen by manual inspection. All recordings were analysed in Praat (Boersma and Weenink 2019) to check whether certain sounds had been mispronounced or omitted. Recordings with such errors and recordings with laughter, stutter, or vocal fry were dismissed. From the best recordings, the pseudoword audio stimuli were then extracted.

2.1.2 Visual stimuli

Visual stimuli were used in the forced-choice task for size rating elicitation and in the cuteness judgement task. Images from van de Vijver and Baer-Henney (2014)[4] were used as visual stimuli in the forced-choice task for size rating elicitation, and each of the pseudoword recordings was randomly matched with one of the visual stimuli for each participant. The visual stimuli from this study depict fantasy creatures which one may best describe as friendly, alien-like beings. In total, all 81 images offered by the study by van de Vijver and Baer-Henney were adopted for the present experiment. All visual stimuli were also used in the cuteness judgement task.

2.2 Procedure

The online experiment consisted of three parts in total. Participants were told to either wear headphones during the experiment or to make sure that their environment was as quiet and as little distracting as possible.

2.2.1 Part 1: Forced-choice task for size rating elicitation

The first part of the experiment consisted of the forced-choice task for size rating elicitation. In this task, pseudoword recordings were randomly matched with one of the visual stimuli. As there were more audio stimuli than visual stimuli (92 vs 81), for each participant, some visual stimuli appeared in two trials. In each

recorded our stimuli produced /ɛː/ and /eː/ and did not merge them. Second, even for participants who do merge both vowels, hearing /ɛː/ as part of a stimulus should come with little to no distraction, as /ɛː/ and /eː/ and their merger are not perceived as strong dialectal markers (Kiesewalter 2019).

4 All images are available from Dinah Baer-Henney upon request.

trial, five differently sized versions of a visual stimulus were presented, with the biggest version on the right and the smallest version on the left side of the screen (Figure 1). For all fantasy creatures, the highest point of their body was close to the upper edge of their image file, and the lowest point of their body was close to the lower edge of their image file. Some of the creatures have a rather tall, thin stature, while others display a smaller, broad stature. The image files of all tall creatures were re-sized to have a height of 200 px, and the image files of all small creatures were re-sized to have a width of 200 px. The width of the tall creatures' and the height of the small creatures' image files varied but were below 200 px. The 200 px version of all image files was used as the smallest visual stimulus. As for the bigger versions within one trial, the second smallest image was two times the size of the smallest; the third biggest image was three times the size of the smallest image; the second biggest image was four times the size of the smallest image; and the biggest image was five times the size of the smallest image. Simultaneously with the display of the five visual stimulus versions, an audio stimulus was played. Participants were told to intuitively decide which of the five visual stimulus versions fitted the audio stimulus best by key-pressing the pertinent visual stimulus' number on their keyboard as quickly as possible.

2.2.2 Part 2: Cuteness judgement task

The second part of the experiment consisted of the cuteness judgement task, which took place after the forced-choice task so as not to prime participants for cuteness. In this task, each trial consisted of one of the 81 visual stimuli. All visual stimuli were shown with the same relative size, that is with the medium size of the forced-choice task. Participants were asked to rate the cuteness of each visual stimulus on a five-point Likert scale, ranging from 1 *nicht niedlich* 'not cute' to 5 *sehr niedlich* 'very cute' (Figure 2).

2.2.3 Part 3: Personal questionnaire

The third and final part of the experiment consisted of a brief questionnaire in which participants were asked to provide their age, their L1/L1s, and their L2/L2s. Participants were recruited by word-of-mouth by all authors and the students of the 'Sound Symbolism' course at Heinrich Heine University Düsseldorf during the winter term 2022–2023.

Overall, 124 participants took part in the experiment; 109 (88%) participants named German as their L1, while in total, 26 different languages were provided by participants as either L1 or L2. Table 2 gives an overview of all languages specified as well as the number of speakers for each language within our group of participants.

As for age, our youngest participant was 17 years old, while our oldest participant was 74 years old. The mean age was 28.9 years with a standard deviation of 11.7, and the median age was 24. Figure 3 displays the distribution of age within our data.

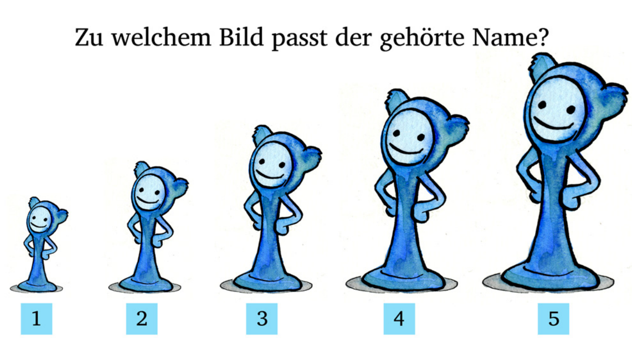

Figure 1: Visual stimulus and response option display during the forced-choice task. The given question is 'Which picture does the name you heard match?'.

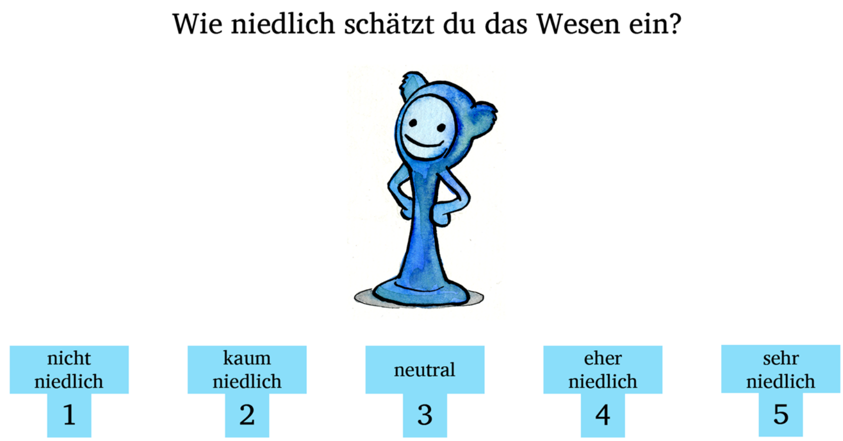

Figure 2: Visual stimulus and response options display during the cuteness judgement task. The given question is 'How cute do you think the creature is?'; the options from 1 to 5 say: 'not cute', 'hardly cute', 'neutral', 'rather cute', and 'very cute'.

Table 2: L1 and L2 data of all participants

Language	L1	L2	Language	L1	L2	Language	L1	L2
Arabic	1	2	Greek	3		Portuguese	1	1
Azerbaijani	3		Japanese	1	4	Rumanian	1	
Chinese*	1	1	Korean		1	Russian	5	3
Croatian		1	Kurdish	1		Spanish		15
Danish		1	Italian		4	Swedish		3
Dutch	1	1	Mandarin		2	Tunisian*	1	
English	5	108	Norwegian		1	Turkish	6	4
French	1	22	Persian	1		Vietnamese	1	
German	109	10	Polish	1		none		8

For languages marked by an asterisk, more specific information is not available.

2.3 Data pre-processing and overview

Ahead of conducting the statistical analysis itself, we checked the sanity of the elicited size rating data. While reaction times were not the focus of the present investigation, they nonetheless provided insight into whether participants truly chose size intuitively during the first part of the experiment (Baayen and Milin 2010). That is,

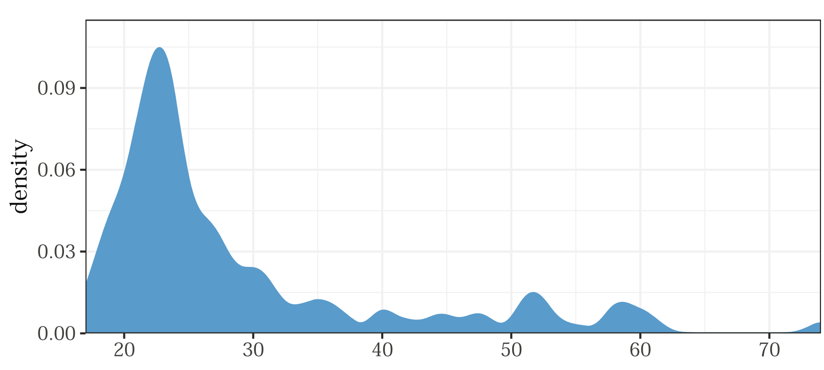

Figure 3: Distribution of age within the data set.

Table 3: Summary of the dependent variable and the categorical and numerical predictors

Dependent variable					
SIZE	1: 832	2: 1,968	3: 2,897	4: 2,833	5: 1,443
Categorical predictors					
VOWEL	/a:/: 1,265	/ɛ:/: 1,283	/e:/: 1,255	/i:/: 1,291	/o:/: 1,073
	/ø:/: 1,277	/u:/: 1,261	/y:/: 1,268		
C1	/d/: 1,610	/f/: 1,674	/j/: 1,583	/k/: 1,699	/ʁ/: 3,407
C2	/d/: 1,602	/f/: 1,659	/j/: 1,589	/k/: 1,715	/ʁ/: 3,408
L1	also German: 2,167	German: 6,862	not German: 944		
L2	English: 5,575	none: 566	also English: 2,544	also German: 944	other: 344
PARTICIPANT	levels: 121				
ITEM	levels: 94				
Numerical predictors	**Median**	**Mean**	**St. Dev.**	**Min**	**Max**
CUTENESS	3.00	2.85	1.26	1.00	5.00
PND	0.00	0.80	1.71	0.00	7.00
AGE	24.00	28.93	11.69	17.00	74.00

data points with reaction times below a threshold of 400 ms were removed, because one may assume that any decisions made in a shorter period of time are made before proper processing (Kutas and Federmeier 2010, Hohlfeld et al. 2015). Data points with reaction times above a threshold of 8,000 ms were also removed. As participants were instructed to decide intuitively, we assumed that everything beyond 20 times the minimum of required time, i.e. 20 × 400 ms, no longer fell into the area of intuitive thinking, but instead is part of analytic processing leading to logical decisions (dual-process theories; Evans and Curtis-Holmes 2005). This resulted in a loss of 1,426 (12.3%) data points, leaving us with 10,188 data points. Finally, the data for the stimuli /do:do:/ and /jo:jo:/ were excluded due to them being homophonous with real words. This led to a loss of a further 215 data points, leaving us with 9,973 data points for the following analyses.

2.4 Statistical analysis

All analyses were conducted in R (R Core Team 2021). The mgcv package (Wood 2017) was used for regression analyses, and the packages ggplot2 (Wickham 2016) and visreg (Breheny and Burchett 2017) for visualisation.

To predict SIZE ratings, an ordinal variable, ordinal logistic regression in generalised additive mixed models was introduced. Generalised additive mixed models were used as they allow non-linear effects. Due to the uncharted territory, our study sets out to explore, i.e. the interaction of cuteness and size ratings from a sound symbolic point of view; we must assume effects of any type. Hence, using modelling approaches which only allow linear effects would be insufficient.

The variables used in the following analyses are either concerned with data directly elicited from participants, with the segmental make-up of the pseudowords, or with biographical information. Following this order, all variables are introduced in the following. Table 3, then, presents an overview of the final data set.

SIZE. This variable contains the size rating for a pertinent trial

CUTENESS. The cuteness judgement of a given participant for a given visual stimulus

VOWEL. The vowel of a pertinent pseudoword

C1/C2. The onset consonant of the first/second syllable of a pertinent pseudoword

PND. Higher phonological neighborhood densities indicate the existence of more similar sounding real words, which in turn may result in unwanted effects such as lower recognition times (Vitevitch and Luce 1998). While the items at hand are pseudowords, that is, they cannot be recognised as known words, higher PNDs may nonetheless result in longer comprehension processes. PND describes the number of words differing in one segment from the pseudoword in question (Marian et al. 2012, 3)

ITEM. The pseudoword used in a given trial

AGE. The age of the relevant participant

L1. The L1s of participants were included as categorical variable. Due to the large number of different languages, with most languages only found in very few participants, we decided against a detailed inclusion of L1s. Instead, the levels of L1 are either 'German', 'also German', or 'not German'

L2. The L2s of participants were also included as categorical variable. As for L1, we decided against a detailed inclusion of languages. Instead, the levels of L2 are either 'also German' (if German is one of several L2s), 'English' (if English is the only L2), 'also English' (if English is one of several L2s of which none is German), 'other' (for L2s that are neither German nor English), or 'none' (for participants without L2s). We decided to include '(also) English' as variable levels because a) 108 of the 124 participants provided English as their L2, and b) there is previous research on cuteness associations in English (see the Introduction)

PARTICIPANT. The ID of the relevant participant

As the main predictors of interest, VOWEL, C1, and C2 entered the analysis as parametric main effects as well as in interaction with CUTENESS. While the reasoning to include these interactions is straightforward – it is these potential interactions we are mainly interested in – VOWEL, C1, and C2 are additionally specified as parametric terms to control for their individual main effects. CUTENESS and PND entered the model as smooth terms to allow for non-linear main effects. Parametric terms are similar to fixed effects in linear mixed-effects models and so are smooth terms, with the important difference that the latter allow for non-linear effects of non-categorical variables. The variables ITEM, AGE, L1, L2, and PARTICIPANT were introduced as random smooth terms, similar to random intercepts in linear mixed-effects models. The formula of this initial ordinal regression model was the following:

```
size ~
C1 +
C2 +
vowel +
s(cuteness, bs = "tp", by = vowel, k = 5) +
s(cuteness, bs = "tp", by = C1, k = 5) +
s(cuteness, bs = "tp", by = C2, k = 5) +
s(cuteness, k = 5) +
s(pnd, k = 7) +
s(age, bs = "re") +
s(L1, bs = "re") +
s(L2, bs = "re") +
s(item, bs = "re") +
s(participant, bs = "re")
```

During the last years, researchers have become more aware of the issue of collinearity when fitting regression models (e.g. Tomaschek et al. 2018). In generalised additive models, a similar issue arises by another name: concurvity (Ramsay et al. 2003). Concurvity occurs when a smooth term in a generalised additive model could be approximated by one or more of the other smooth terms in the same model. Just like collinearity, concurvity causes issues of reliability, as estimates become unstable, and with that issues of interpretability. To avoid such issues, the initial model was checked for concurvity issues. The check revealed that multiple variable combinations lead to concurvity in the initial model: PND and ITEM, AGE and PARTICIPANT, L1 and PARTICIPANT, and L2 and PARTICIPANT. From a conceptual perspective, these variable pairings are not surprising. PND is measured via the segmental make-up of a pseudoword. Hence, both variables PND and ITEM should be closely related, and one can easily predict one with the other. For PARTICIPANT, on the one hand, and AGE, L1, and L2, on the other hand, we can assume a similar situation. Therefore, another model was fitted with similar specifications as given above, but without ITEM and PARTICIPANT random smooth terms. The decision to exclude ITEM and PARTICIPANT instead of their related variables is based on what the variables are able to tell us, should they turn out to show significant effects. The excluded two cannot really tell us anything as their scope is too broad. The variables retained, however, have a narrower scope (L1 and L2) or are clearly restricted in terms of interpretation (AGE). The reduced model did not show any concurvity issues; its results are presented in the following section.

3 Results

The R script and data used for the following analyses can be found at the OSF: https://osf.io/qck9r/.

3.1 The cuteness of alien creatures

Before we dive into the potential sound symbolic interactions of cuteness and size, we provide insight into the data elicited by the cuteness judgement task. Overall, alien creatures were rated with a mean cuteness of 2.85 with a mean standard deviation of 1.17. The cutest alien was rated with a mean cuteness of 3.98 (SD: 1.11), while the least cute alien was rated with a mean cuteness of 1.88 (SD: 1.08). The cutest and the least cute alien are shown in Figure 4. Figure 5 presents the mean cuteness ratings and their standard deviations for all alien creatures.

3.2 Effects on size ratings

The results of the ordinal regression model are given in Tables 4 and 5. For the parametric terms, we provide the β estimates and the corresponding standard errors (SE), *z*-values, and *p*-values. For the smooth terms, the estimated degrees of freedom, the reference degrees of freedom, the χ^2 values, and the *p*-values are given.

3.2.1 Individual effects of vowels, consonants, and cuteness

Let us first take a look at the main effect of vowels, consonants, and cuteness. Bonferroni-corrected pairwise Tukey comparisons for VOWEL based on the predictions of the ordinal regression model are given in Table 6.

Figure 6 is a conditional plot showing the effect of VOWEL holding all other variables constant (median for numeric variables, most common level for categorical variables). All following effect plots are also conditional plots. It can be seen that /aː/, /oː/, and /øː/ come with the largest estimated mean sizes, while /iː/ comes with the smallest estimated mean size.

For C1 and C2, Bonferroni-corrected pairwise Tukey comparisons based on the predictions of the ordinal regression model found the significant differences shown in Table 7. For the onset consonant of the first

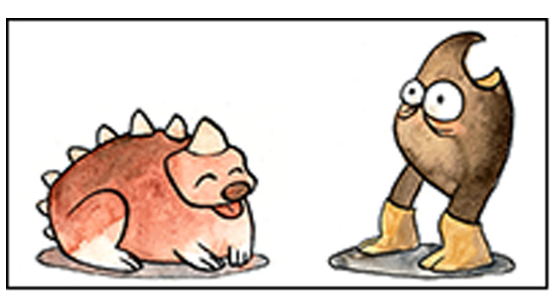

Figure 4: The alien creatures rated cutest (left) and least cute (right).

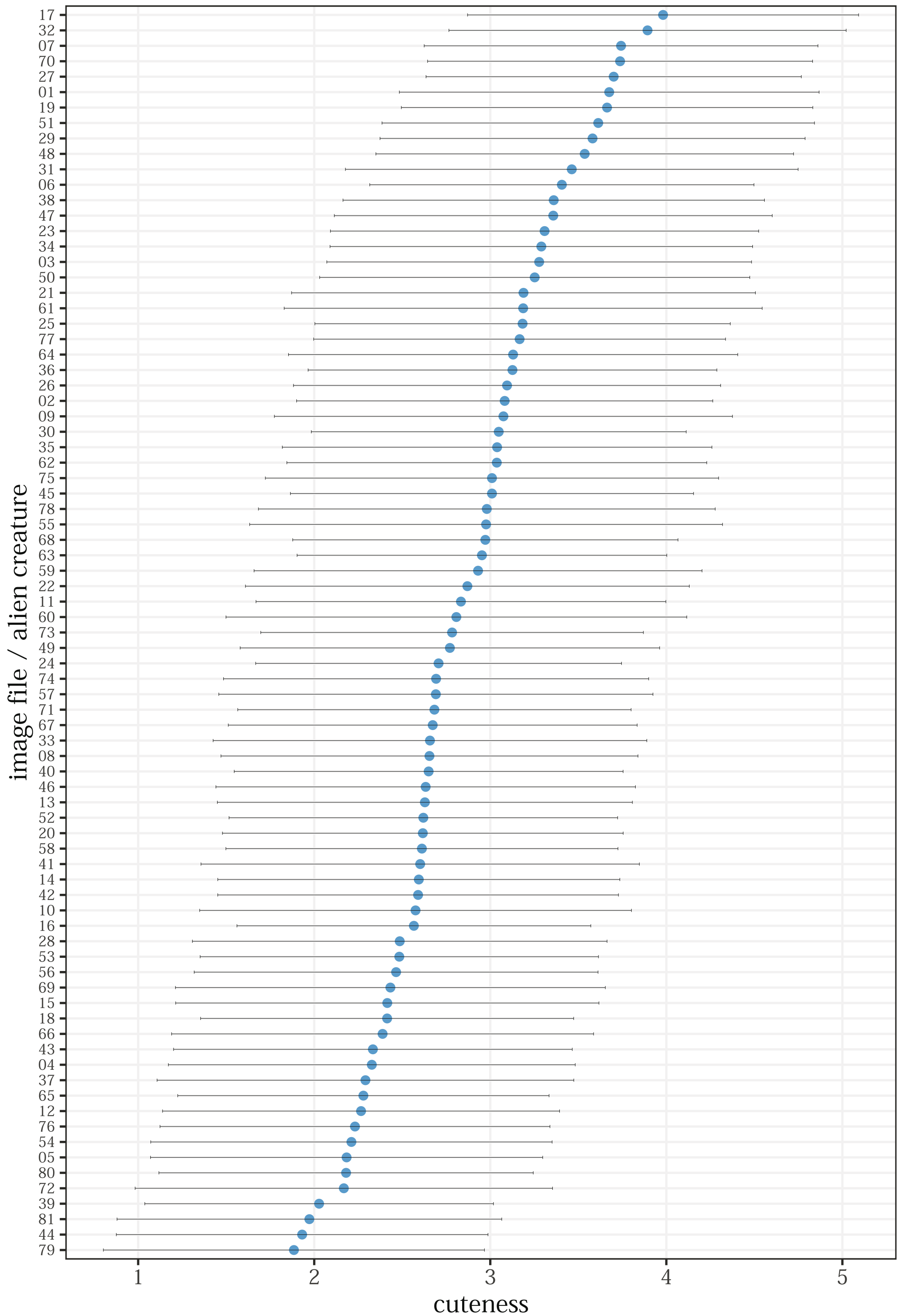

Figure 5: Cuteness ratings for all 81 alien creatures. Dots represent the mean, whiskers display ±1 standard deviation.

Table 4: Summary of the parametric terms in the ordinal regression model

Parametric terms	β Estimate	SE	*z*-Value	*p*-Value
(Intercept)	2.103	0.134	15.679	>0.001
C1f	−0.086	0.073	−1.185	0.236
C1j	−0.153	0.073	−2.099	0.036
C1k	−0.083	0.074	−1.113	0.266
C1r	0.199	0.065	3.049	0.002
C2f	−0.002	0.073	−0.021	0.983
C2j	0.009	0.073	0.129	0.897
C2k	0.086	0.073	1.170	0.242
C2r	0.313	0.066	4.739	>0.001
VOWELɛ	−0.627	0.072	−8.731	>0.001
VOWELe	−0.712	0.072	−9.835	>0.001
VOWELi	−1.814	0.097	−18.712	>0.001
VOWELo	−0.376	0.087	−4.343	>0.001
VOWELø	−0.325	0.073	−4.470	>0.001
VOWELu	−0.681	0.077	−8.810	>0.001
VOWELy	−0.851	0.075	−11.409	>0.001

Table 5: Summary of the smooth terms in the ordinal regression model

Smooth terms	edf	Ref.df	Chi.sq	*p*-Value
CUTENESS	1.002	1.003	0.004	0.966
PND	2.750	3.376	8.675	0.036
AGE	0.449	1.000	1.100	0.177
L1	1.409	2.000	65.173	0.028
L2	2.980	4.000	14.347	0.336
Interactions	**edf**	**Ref.df**	**Chi.sq**	***p*-value**
CUTENESS:VOWELa	2.783	3.320	25.636	>0.001
CUTENESS:VOWELɛ	1.191	1.355	0.274	0.846
CUTENESS:VOWELe	1.004	1.007	0.049	0.830
CUTENESS:VOWELi	2.329	2.851	18.715	>0.001
CUTENESS:VOWELo	1.035	1.513	2.410	0.333
CUTENESS:VOWELø	1.890	2.336	2.038	0.398
CUTENESS:VOWELu	1.003	1.006	0.183	0.670
CUTENESS:VOWELy	1.625	2.007	3.151	0.208
CUTENESS:C1d	1.001	1.001	0.364	0.547
CUTENESS:C1f	1.001	1.002	0.662	0.417
CUTENESS:C1j	1.001	1.002	0.057	0.813
CUTENESS:C1k	0.002	0.003	0.000	0.988
CUTENESS:C1r	1.003	1.006	0.441	0.511
CUTENESS:C2d	1.222	1.719	1.679	0.394
CUTENESS:C2f	1.001	1.001	0.000	1.000
CUTENESS:C2j	1.001	1.001	1.083	0.298
CUTENESS:C2k	1.719	2.213	1.875	0.409
CUTENESS:C2r	1.001	1.003	0.643	0.424

Table 6: Bonferroni-corrected pairwise Tukey comparisons for all levels of VOWEL

	a:	ɛ:	e:	i:	o:	ø:	u:
ɛ:	***						
e:	***	n.s.					
i:	***	***	***				
o:	n.s.	*	***	***			
ø:	n.s.	n.s.	**	***	n.s.		
u:	***	n.s.	n.s.	***	***	*	
y:	***	*	n.s.	***	***	***	n.s.

Significance codes: '***' $p < 0.001$, '**' $p < 0.01$, '*' $p < 0.05$.

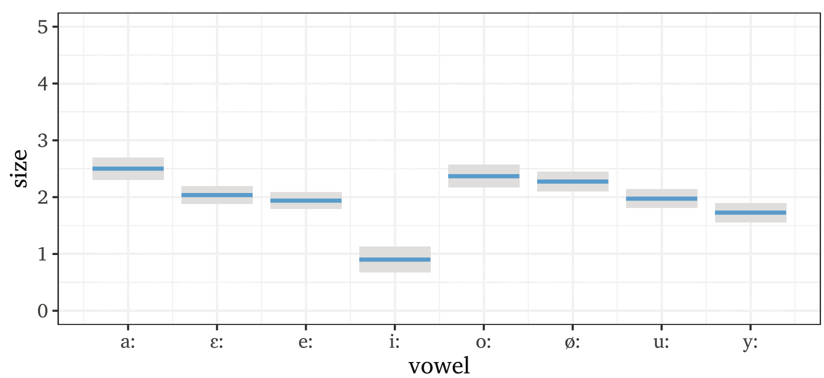

Figure 6: Effect of VOWEL in the ordinal regression model.

syllable, significant differences are found comparing /ʁ/ to /d, f, j/, and /k/. For the onset consonant of the second syllable, significant differences are found comparing /ʁ/ to /d, f/, and /j/.

Figure 7 presents the estimated differences between the individual C1 and C2 consonants. Overall, the differences between consonants appear to be smaller than the differences between vowels. For C1, /ʁ/ comes with the biggest estimated mean size, while the estimates of /f/ and /j/ are significantly smaller. For C2, /ʁ/ also comes with the biggest estimated size ratings, being significantly bigger than /d, f/, and /j/.

Finally, CUTENESS on its own did not reach significance.

3.2.2 Interactions of vowels, consonants, and cuteness

Let us now move on to the interactions of interest: CUTENESS and VOWEL, CUTENESS and C1, and CUTENESS and C2. Figure 8 presents the effects of the interaction of CUTENESS and VOWEL, which reached significance for /iː/ and /aː/.

Table 7: Bonferroni-corrected pairwise Tukey comparisons for all levels of C1 and C2

	C1				C2			
	d	f	j	k	d	f	j	k
f	n.s.				n.s.			
j	n.s.	n.s.			n.s.	n.s.		
k	n.s.	n.s.	n.s.		n.s.	n.s.	n.s.	
ʁ	*	***	***	***	***	***	***	n.s.

Significance codes: '***' $p < 0.001$, '**' $p < 0.01$, '*' $p < 0.05$.

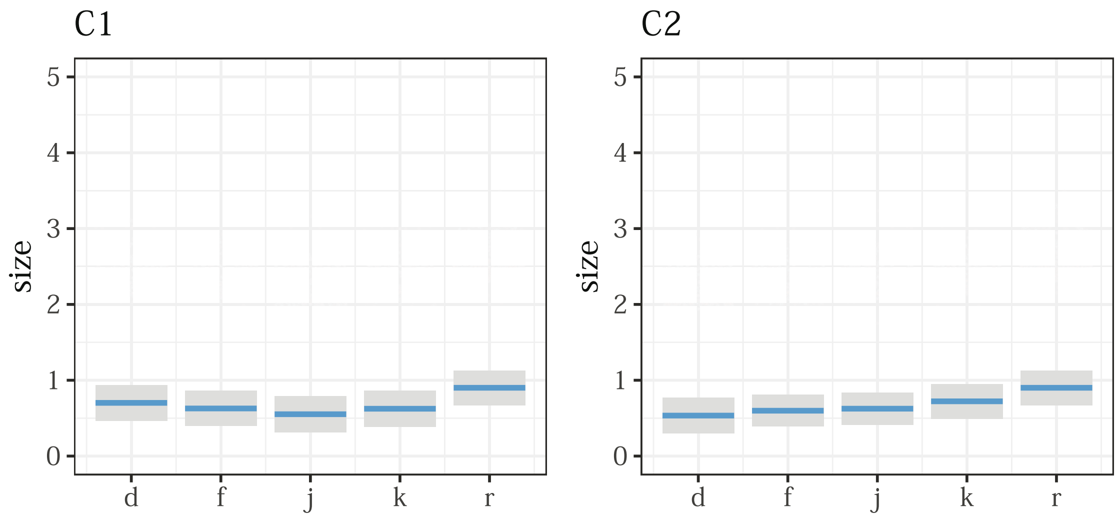

Figure 7: Effects of C1 (left plot) and C2 (right plot) in the ordinal regression model.

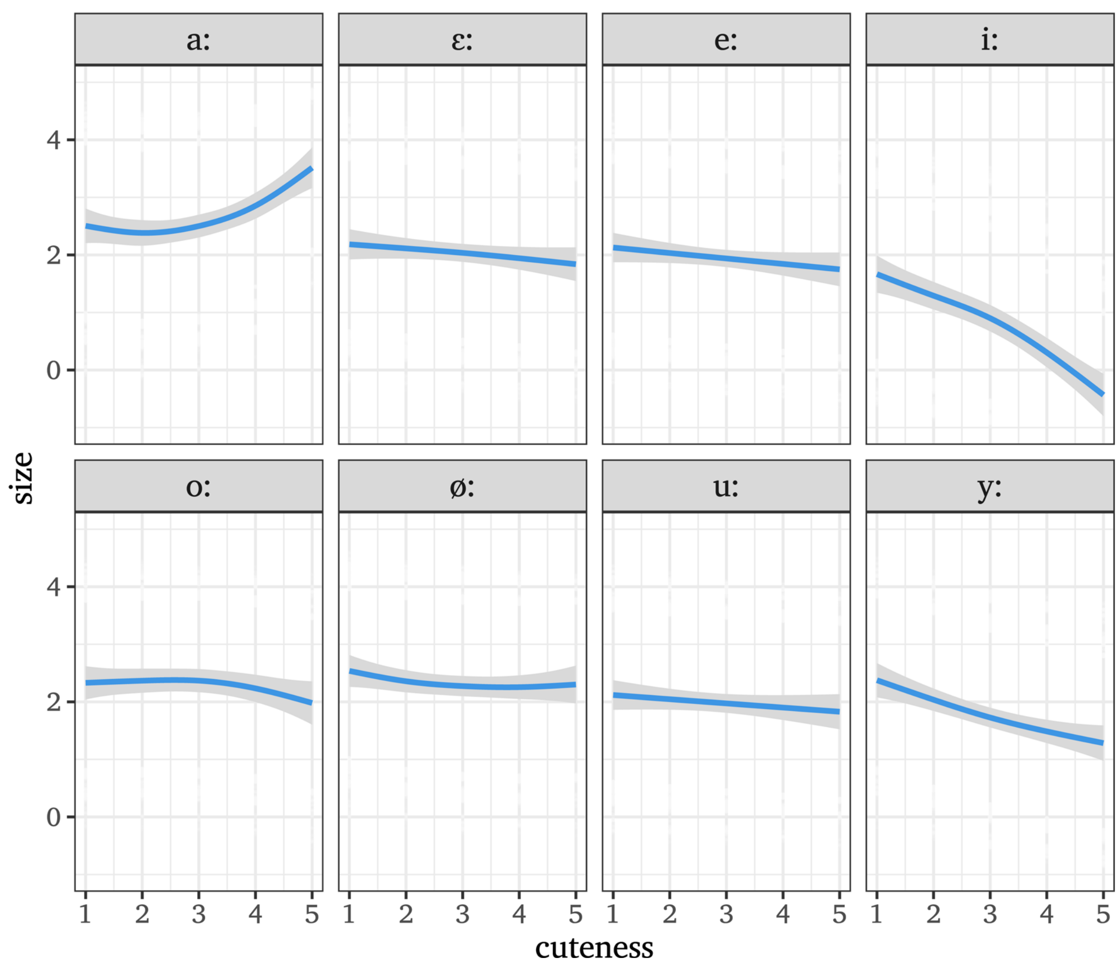

Figure 8: Effects of the interaction of CUTENESS and VOWEL in the ordinal regression model. Each panel displays the effect found for a given vowel.

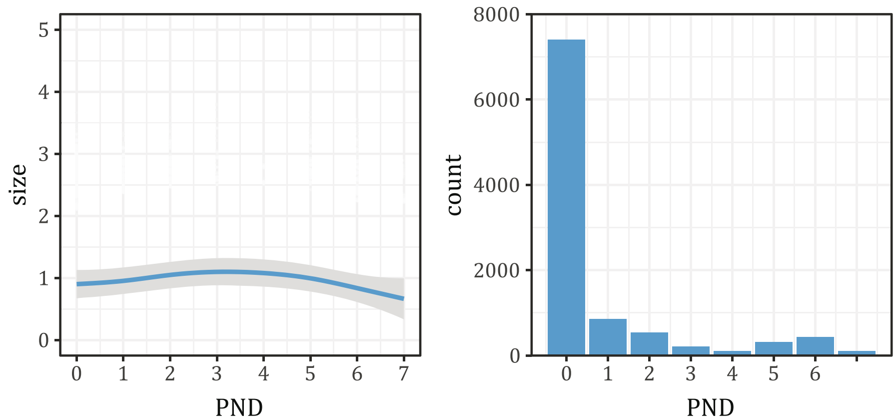

Figure 9: Effect of PND (left plot) in the ordinal regression model and distribution of PND (right plot) in the data.

If an alien's pseudoword name contains /i:/ as nuclei, cuter aliens are judged to be smaller. However, if the pseudoword contains /a:/ as nuclei, cuter aliens are taken to be bigger.

The interactions of CUTENESS and C1 and C2 did not reach significance.

3.2.3 Effects of PND, age, L1, and L2

Finally, let us have a look at the results for the remaining variables. The effect of PND is shown in Figure 9, alongside the distribution of PND in the data. Taking both the effect and the distribution into account, we conclude that in trials with pseudowords with higher PND values, as for example for /ju:ru:/ with six neighbours (*Guru* /gu:ʁu/, *Yuri* /ju:ʁi:/, *Jury* /ʒy:ʁi:/, *Jura* /ju:ʁa:/, *Juri* /ʒu:ʁi:/, *Juhu* /ju:hu:/), smaller sizes were chosen by participants than in trials with pseudowords with smaller PND values, as for example for /fo:fo:/ with one neighbour (*Foto* /fo:to/).

Bonferroni-corrected pairwise Tukey comparisons for L1 based on the predictions of the ordinal regression model, as shown in Table 8, reveal that there is a difference between those participants who speak German as sole L1 and those who have additional L1s. The effect of L1 is shown in Figure 10. That is, monolingual L1 speakers of German come with higher estimated size ratings.

Neither the effect of L2 nor the effect of AGE reached significance.

4 Discussion and conclusion

The present article set out to answer the question of whether size and cuteness associations interact. Using a forced-choice task, size ratings were elicited for combinations of pseudowords and visual stimuli. Disyllabic pseudowords were constructed to contain only one of eight German long vowels and the same or different onset consonants in either syllable. Cuteness judgements for all visual stimuli were elicited on a five-point

Table 8: Bonferroni-corrected pairwise Tukey comparisons for all levels of L1

	also German	**German**
German	***	
no German	n.s.	n.s.

Significance codes: '***' $p < 0.001$, '**' $p < 0.01$, '*' $p < 0.05$.

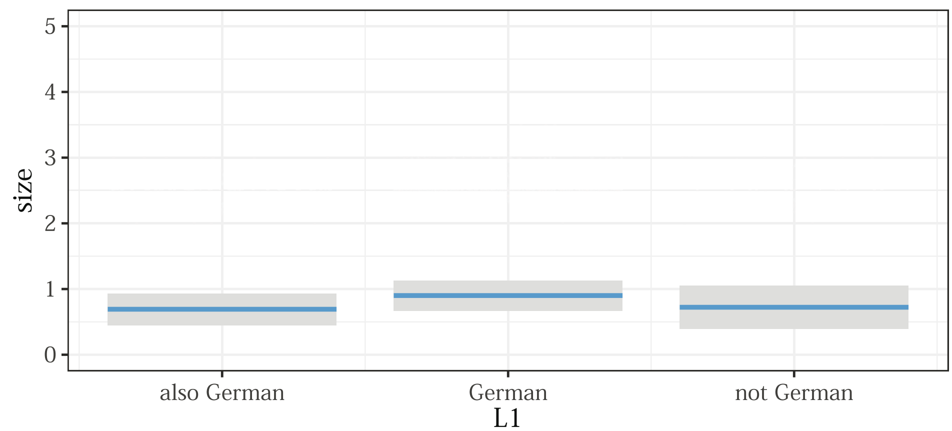

Figure 10: Effect of L1 in the ordinal regression model.

Likert scale. Participants made use of the entire Likert scale judging the cuteness of alien creatures, that is, the visual stimuli employed in the experiment. While there were differences in the cuteness judgements between participants, overall, all cuteness ratings show roughly the same amount of deviation. Across all participants, a number of visual stimuli stood out as particularly cute and least cute.

Making use of these cuteness judgements, ordinal logistic regression with SIZE ratings as dependent variable was fitted in a generalised additive mixed model. Within the model, the interactions of interest were contained: a) CUTENESS and VOWEL, b) CUTENESS and C1 (the onset consonant of the first syllable), and c) CUTENESS and C2 (the onset consonant of the second syllable). Alongside the interactions, the main effects of the variables included in the interactions were given, as well as several potentially confounding variables: PND, participants' AGE, L1s, and L2s.

Among the potentially confounding variables, only PND and L1 showed significant effects. Participants judged the size of auditory stimuli with a higher number of phonological neighbours to be generally smaller, while auditory stimuli with one to no neighbours were rated as bigger. A potential explanation for this finding might lie in the semantics of phonological neighbours for those stimuli that come with such neighbours. Following the idea of linear discriminative learning (e.g. Baayen et al. 2019), one assumes that phonology and semantics are not separable but tightly connected for entries in the mental lexicon. That is, if for a stimulus phonological neighbours are activated, so are their semantics. Analogously, if a stimulus does not activate phonological neighbours, the semantics of other entries are not activated either. Thus, semantically, the idea of 'alien creature' is more prominent for a stimulus without neighbours than for stimuli with activated neighbours, whose semantics in turn are co-activated and, therefore, render the idea of 'alien creature' less prominent. With alien creatures being mostly represented as at least human size in popular culture, the idea of an 'alien creature' may thus evoke higher size ratings when semantically activated without competitors. The effect found for phonological neighbourhood size makes a very interesting case for future research, especially for research making use of linear discriminative learning. As for L1s, it was found that monolingual L1 speakers of German generally rate size to be larger than multilingual L1 speakers of German. While L2 as a variable did not reach significance, most participants provided at least one of eighteen languages as L2. Due to the rather large number of combinations of different languages as L1s and L2s in our data, we refrain from any speculations as to why the effect found for L1 exists.

Moving on to the individual effects of VOWEL, C1, C2, and CUTENESS, the ordinal regression model revealed that CUTENESS did not reach significance. For the eight vowels under investigation, it was found that /a:/ comes with the highest size ratings and that /i:/ comes with the lowest size ratings. These findings are in line with previous research on size sound symbolism in vowels (e.g. Berlin 2006, Blasi et al. 2016, Johansson 2017, Shinohara and Kawahara 2010, Westbury et al. 2018). However, /a/ was not alone with its highest rated size: /o:/ and /ø:/ were not rated significantly smaller. A potential explanation may be that /a:/, /o:/, and /ø:/ are central or back vowels, while /iː/ is a front vowel. That is, size associations of vowels might not only be related to values of the first formant, i.e. the higher the bigger, but also with values of the second formant, i.e. the lower the bigger, or a combination of both. However, further research is needed to corroborate this idea. For the five consonants under investigation, we found that /ʁ/ is associated with bigger size than /d, f, j/, and /k/ in the onset of the first syllable and with bigger size than /d, f/, and /j/ in the onset of the second syllable. These findings are somewhat surprising as previous research reported that size is associated with voiced obstruents (Newman 1933, Hamano 1986, Shinohara and Kawahara 2010, Kawahara and Shinohara 2012). Hence, one would assume that /d/ should come with size ratings just as high as /ʁ/ for both onsets. Notably, the size ratings for /ʁ/ are not significantly different than those for /k/ for the onset of the second syllable. As both sounds are produced dorsally, their shared tongue gesture may be an explanation. Further, uvular /ʁ/ is produced furthest back, followed by velar /k/, which in turn is followed by palatal /j/, alveolar /d/, and labiodental /f/. Hence, it seems as if not only voiced obstruents are associated with larger size but also obstruents produced further back in the oral cavity. One potential explanation for the divergent findings for /ʁ/ might lie in what Ćwiek defined as 'sound-symbolic functional load', the distribution of sound-symbolic communicative force through available means (Ćwiek 2022, 57). The sound /ʁ/ is rather rare across languages (Moran and McCloy 2019), and the studies cited thus far are concerned with languages that lack /ʁ/. In turn, these languages cannot associate largeness with /ʁ/. German, however, knows /ʁ/ and, apparently, it is a good candidate sound to be associated with largeness, probably due to its distinctiveness (Anselme et al. 2023).

Let us now take a closer look at the interactions of CUTENESS and VOWEL, CUTENESS and C1, and CUTENESS and C2. While the interactions of both onset consonants and CUTENESS did not reach significance, the interaction of CUTENESS and VOWEL did. More precisely, the interaction showed significant effects for two vowels: /aː/ and /iː/. For the open vowel /aː/, it was found that the cuter the alien creature described by a pertinent pseudoword, the bigger its associated size. For the close vowel /iː/, it was found that the cuter the alien creature described by a pertinent pseudoword, the smaller its associated size. For all remaining vowels, no significant effect was found. Previous research on size sound symbolism found that /aː/ is the 'biggest', while /iː/ is the 'smallest' vowel (e.g. Berlin 2006, Blasi et al. 2016, Johansson 2017, Shinohara and Kawahara 2010, Westbury et al. 2018). Interestingly, the 'size' of /aː/ is judged to be even bigger with increasing cuteness; that is the 'biggest' vowel becomes even bigger. Analogously, the 'size' of /iː/ is judged to be even smaller with increasing cuteness, i.e. the 'smallest' vowel becomes even smaller. It appears that cuteness modifies size sound symbolism in vowels at its extremes.

How can we explain this novel finding? One approach to an explanation lies in the findings of Klink (2000), who found that positive traits are associated with front vowels in American English. If we assume that cuteness is a positive trait, a connection of the front vowel /iː/ and cuteness is straightforward. This connection, then, might potentially enhance the size sound symbolic effect of /iː/ in cute contexts. However, this approach does not explain why /aː/, an open central unrounded vowel, would be influenced by cuteness in its size sound symbolic effect and, especially, why the nature of this effect should be contrary to the one found for /iː/, a close front unrounded vowel. One might assume that Klink's findings only hold for American English and that speakers of German also associate central vowels with positive traits. However, this would not account for the directionality of the effect itself.

Further, similar to Kumagai (2020), one might argue that the German word used in the cuteness judgement task and its different meanings potentially confounded the interaction of cuteness judgements and vowels. The attested two meanings of *niedlich* 'cute' are 1) appealing graceful, childlike pretty, nice (and lively) and 2) little DWDS (2023). Hence, even if the less frequent meaning, i.e. meaning 2), was 'in use' by our participants, cuteness and smallness would still go hand in hand. In turn, an effect should be found for /iː/, the 'smallest' vowel, but not for /aː/.

Finally, the infant schema might deliver an explanation. For /iː/, a connection to the infant schema is straightforward. As babbling infants mostly produce high-frequency vowel-like sounds, humans associate high-frequency vowels such as /iː/ with infants, i.e. with cuteness. For /aː/, however, a connection to the infant schema is less straightforward. One might assume that the proportionally big heads and eyes of infants are the meaningful factor linking cuteness and largeness. Still, this connection would not explain why with increasing cuteness the associated size of /iː/ decreases, while the associated size of /aː/ increases. One potential explanation might be the situation in which the relevant vowels are produced by infants. While /iː/ is one of the babbling sounds and as such part of a group of positive inviting cues (Kringelbach et al. 2016), /aː/, on the contrary, is articulatorily close to the most frequent vowel in infants' crying, /æ/ (Irwin and Curry 1941), and, thus, part of a group of negative aversive cues (Kringelbach et al. 2016). Crying is typically perceived as loud and listeners generally relate high amplitudes to largeness (Grassi et al. 2013). In other words, the difference in cue nature and the association with largeness might explain the effect of cuteness on size associations for /aː/: /aː/-like sounds produced by cute entities such as babies are associated with loudness, which in turn is associated with largeness. Thus, alien creatures whose names contain /aː/ as nuclei are associated with increasing size the cuter the alien is. /oː/ and /øː/, the other two vowels that were among the highest rated vowels are not /aː/-like sounds, and therefore, no significant interaction of cuteness and these two vowels was found.

The apparent influence of one sensory information type on another's sound symbolic effect comes with threefold consequences. As a warning, this finding cautions against focussing on but one type of sensory information in sound symbolism research, i.e. at least in contexts in which more than one type of sensory information is present or applicable. As a task, it calls for reanalyses of previous research on sound symbolism in which more than one type of sensory information was present but only one was considered in analyses. As an opportunity, the present findings shine a light on a potentially wide variety of different types of sensory information that might influence each other in their sound symbolic effects. Investigating such intricate interactions will further inform not only sound symbolism research but also linguistics as a whole.

In sum, the present study reported novel findings on size sound symbolism in vowels in that it is modulated by cuteness. It appears that sound symbolic effects may manifest in intricate interactions when multiple visual dimensions such as size and cuteness are considered. This finding calls for the incorporation of multiple types of sensory information where and whenever applicable in research on sound symbolism.

Acknowledgements: The authors would like to thank the students of the course 'Sound Symbolism', Janina Esser for providing audio stimuli, and the members of the Department of English Language and Linguistics at Heinrich-Heine-Universität Düsseldorf, with special thanks to Viktoria Schneider who inspired parts of this research. The authors appreciate the feedback provided by the audiences of the 16th International Cognitive Linguistics Conference and the 5th Phonetics and Phonology in Europe conference and are most grateful for the feedback and suggestions by two anonymous reviewers.

Funding information: The position of DS was funded by the DFG project "The semantics of derivational morphology" (Grant PL 151/11-1 to Ingo Plag) at Heinrich-Heine-Universität Düsseldorf. The publication of this article was supported by the HHU Open Access Fund.

Author contributions: DS designed the study. All authors recruited participants and were involved in the modelling and statistical analyses. DS wrote the first draft of the manuscript; all authors were involved in revising the manuscript. The submitted version was read and approved by all authors.

Conflict of interest: The authors state no conflict of interest.

Data availability statement: The data sets analysed during the current study are available via the OSF repository, https://osf.io/qck9r/.

References

Ahlner, Felix and Jordan Zlatev. 2010. "Cross-modal iconicity: A cognitive semiotic approach to sound symbolism." *Sign Systems Studies* 38(1/4), 298–348. doi: 10.12697/SSS.2010.38.1-4.11.

Anselme, Rémi, François Pellegrino, and Dan Dediu. 2023. "What's in the r? A review of the usage of the r symbol in the Illustrations of the IPA." *Journal of the International Phonetic Association* 17(5), 1–30. doi: 10.1017/S0025100322000238.

Baayen, R. Harald, Yu-Ying Chuang, Elnaz Shafaei-Bajestan, and James P. Blevins. 2019. “The discriminative lexicon: A unified computational model for the lexicon and lexical processing in comprehension and production grounded not in (de)composition but in linear discriminative learning.” *Complexity* 2019, 4895891. doi: 10.1155/2019/4895891.

Baayen, R. Harald and Petar Milin. 2010. “Analyzing reaction times.” *International Journal of Psychological Research* 3(2), 12–28. doi: 10.21500/20112084.807.

Berlin, Brent. 1995. “Evidence for pervasive synesthetic sound symbolism in ethnozoological nomenclature.” In *Sound symbolism*, edited by Leanne Hinton, Johanna Nichols, and John J. Ohalla, p. 76–93. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511751806.006.

Berlin, Brent. 2006. “The first congress of ethnozoological nomenclature.” *Journal of the Royal Anthropological Institute* 12(s1), S23–44. doi: 10.1111/j.1467-9655.2006.00271.x.

Blasi, Damián E., Søren Wichmann, Harald Hammarström, Peter F. Stadler, and Morten H. Christiansen. 2016. “Sound-meaning association biases evidenced across thousands of languages.” *Proceedings of the National Academy of Sciences of the United States of America* 113(39), 10818–23. doi: 10.1073/PNAS.1605782113.

Boersma, Paul and David Weenink. 2019. *Praat: Doing phonetics by computer*. http://www.praat.org/.

Breheny, Patrick and Woodrow Burchett. 2017. “Visualization of regression models using visreg.” *The R Journal* 9(2), 56. doi: 10.32614/RJ-2017-046.

Bremner, Andrew J., Serge Caparos, Jules Davidoff, Jan de Fockert, Karina J. Linnell, and Charles Spence. 2013. ““Bouba” and “Kiki” in Namibia? A remote culture make similar shape–sound matches, but different shape–taste matches to Westerners.” *Cognition* 126(2), 165–72. doi: 10.1016/J.COGNITION.2012.09.007.

Caselli, Naomi K., Michael K. Caselli, and Ariel M. Cohen-Goldberg. 2016. “Inflected words in production: Evidence for a morphologically rich lexicon.” *Quarterly Journal of Experimental Psychology* 69(3), 432–54. doi: 10.1080/17470218.2015.1054847.

Ćwiek, Aleksandra. 2022. “Iconicity in language and speech.” *Humboldt-Universität zu Berlin dissertation*. doi: 10.18452/24544.

Ćwiek, Aleksandra, Susanne Fuchs, Christoph Draxler, Eva Liina Asu, Dan Dediu, Katri Hiovain, Shigeto Kawahara, et al. 2022. “The bouba/kiki effect is robust across cultures and writing systems”. *Philosophical Transactions of the Royal Society B: Biological Sciences* 377(1841). doi: 10.1098/rstb.2020.0390.

Diffloth, Gérard. 1994. “I: big, a: small.” In *Sound symbolism*, edited by Leanne Hinton, Johanna Nichols, and John Ohala, p. 107–14. Cambridge: Cambridge University Press.

D’Onofrio, Annette. 2013. “Phonetic detail and dimensionality in sound-shape correspondences: Refining the bouba-kiki paradigm.” *Language and Speech* 57(3), 367–93. doi: 10.1177/0023830913507694.

DWDS. niedlich. In *DWDS – Digitales Wörterbuch der deutschen Sprache*. Retrieved September 16, 2023. https://www.dwds.de/wb/niedlich.

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

Evans, Jonathan St. B. T. and Jodie Curtis-Holmes. 2005. “Rapid responding increases belief bias: Evidence for the dual-process theory of reasoning.” *Thinking & Reasoning* 11(4), 382–9. doi: 10.1080/13546780542000005.

Gahl, Susanne. 2008. “Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech.” *Language* 84(3), 474–96. doi: 10.1353/lan.0.0035.

Grassi, Massimo, Massimiliano Pastore, and Guillaume Lemaitre. 2013. “Looking at the world with your ears: How do we get the size of an object from its sound?” *Acta Psychologica* 143(1), 96–104. doi: 10.1016/j.actpsy.2013.02.005.

Hamano, Shoko. 1986. *The sound-symbolic system of Japanese*. University of Florida, dissertation.

Hohlfeld, Annette, Manuel Martín-Loeches, and Werner Sommer. 2015. “Is semantic processing during sentence reading autonomous or controlled? Evidence from the N400 component in a dual task paradigm.” *Advances in Cognitive Psychology* 11(2), 42. doi: 10.5709/ACP-0170-2.

Irwin, Orvis C., and Thayer Curry. 1941. “Vowel elements in the crying vocalization of infants under ten days of age.” *Child Development* 12(2), 99–109.

Johansson, Niklas. 2017. “Tracking linguistic primitives.” In *Dimensions of iconicity*, edited by Angelika Zirker, Matthias Bauer, Olga Fischer, and Christina Ljungberg, p. 39–62. Amsterdam: John Benjamins. doi: 10.1075/ill.15.03joh.

Kawahara, Shigeto. 2012. “Lyman’s Law is active in loanwords and nonce words: Evidence from naturalness judgment studies.” *Lingua* 122(11), 1193–206. doi: 10.1016/j.lingua.2012.05.008.

Kawahara, Shigeto and Kazuko Shinohara. 2012. “A tripartite trans-modal relationship among sounds, shapes and emotions: A case of abrupt modulation.” In *Proceedings of the Annual Meeting of the Cognitive Science Society*.

Kiesewalter, Carolin. 2019. *Zur subjektiven Dialektalität regiolektaler Aussprachemerkmale des Deutschen*. Stuttgart: Franz Steiner Verlag. doi: 10.25162/9783515124430.

Kilpatrick, Alexander, Aleksandra Ćwiek, Eleanor Lewis, and Shigeto Kawahara. 2023. “A cross-linguistic, sound symbolic relationship between labial consonants, voiced plosives, and Pokémon friendship.” *Frontiers in Psychology* 14, 1113143. doi: 10.3389/fpsyg.2023.1113143.

Kleiner, Stefan. 2011. *Atlas zur Aussprache des deutschen Gebrauchsstandards (AADG). Unter Mitarbeit von Ralf Knöbl*. https://prowiki.ids-mannheim.de/bin/view/AADG/.

Klink, Richard R. 2000. “Creating brand names with meaning: The use of sound symbolism.” *Marketing Letters* 11(1), 5–20. doi: 10.1023/A:1008184423824.

Köhler, Wolfgang. 1929. *Gestalt psychology*. New York: Liveright.

Kringelbach, Morten L., Eloise A. Stark, Catherine Alexander, Marc H. Bornstein, and Alan Stein. 2016. "On cuteness: Unlocking the parental brain and beyond." *Trends in Cognitive Sciences* 20(7), 545–58. doi: 10.1016/j.tics.2016.05.003.

Kumagai, Gakuji. 2019. "A sound-symbolic alternation to express cuteness and the orthographic Lyman's Law in Japanese." *Journal of Japanese Linguistics* 35(1), 39–74. doi: 10.1515/jjl-2019-2004.

Kumagai, Gakuji. 2020. "The pluripotentiality of bilabial consonants: The images of softness and cuteness in Japanese and English." *Open Linguistics* 6(1), 693–707. doi: 10.1515/opli-2020-0040.

Kumagai, Gakuji and Shigeto Kawahara. 2017. "How abstract is sound symbolism? Labiality and diaper names in Japanese [in Japanese]." In *Proceedings of the 31st Phonetic Society of Japan*, p. 49–54.

Kutas, Marta and Kara D. Federmeier. 2010. "Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP)." *Annual Reviews* 62, 621–47. doi: 10.1146/ANNUREV.PSYCH.093008.131123.

Lehmann, Vicky, Elisabeth M. J. Huis in't Veld, and Ad JJM Vingerhoets. 2013. "The human and animal baby schema effect: Correlates of individual differences." *Behavioural Processes* 94, 99–108. doi: 10.1016/j.beproc.2013.01.001.

Lohmann, Arne. 2018. "Time and thyme are not homophones: A closer look at Gahl's work on the lemma-frequency effect, including a reanalysis." *Language* 94(2), e180–90. doi: 10.1353/lan.2018.0032.

Lorenz, Konrad. 1943. "Die angeborenen Formen möglicher Erfahrung." *Zeitschrift für Tierpsychologie* 5(2), 235–409. doi: 10.1111/J.1439-0310.1943.TB00655.X.

MacNeilage, Peter F. and Barbara L. Davis. 2000. "On the origin of internal structure of word forms." *Science* 288(5465), 527–31. doi: 10.1126/SCIENCE.288.5465.527.

Marian, Viorica, James Bartolotti, Sarah Chabal, and Anthony Shook. 2012. "Clearpond: Cross-linguistic easy-access resource for phonological and orthographic neighborhood densities." *PLoS ONE* 7(8), e43230. doi: 10.1371/journal.pone.0043230.

Mathôt, Sebastiaan, Daniel Schreij, and Jan Theeuwes. 2012. "OpenSesame: An open-source, graphical experiment builder for the social sciences." *Behavior Research Methods* 44(2), 314–24. doi: 10.3758/s13428-011-0168-7.

Maurer, Daphne, Thanujeni Pathman, and Catherine J. Mondloch. 2006. "The shape of boubas: Sound–shape correspondences in toddlers and adults." *Developmental Science* 9(3), 316–22. doi: 10.1111/J.1467-7687.2006.00495.X.

Moran, Steven and Daniel McCloy. 2019. *PHOIBLE 2.0*. Jena: Max Planck Institute for the Science of Human History. doi: 10.5281/zenodo.2593234.

Newman, Stanley S. 1933. "Further experiments in phonetic symbolism." *The American Journal of Psychology* 45(1), 53. doi: 10.2307/1414186.

Nielsen, Alan K. S. and Drew Rendall. 2013. "Parsing the role of consonants versus vowels in the classic Takete-Maluma phenomenon." *Canadian Journal of Experimental Psychology* 67(2), 153–63. doi: 10.1037/A0030553.

Perfors, Amy. 2004. "What's in a name? The effect of sound symbolism on perception of facial attractiveness." *Proceedings of the Annual Meeting of the Cognitive Science Society* 26(26), 1617.

R Core Team. 2021. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.r-project.org/.

Ramachandran, Vilayanur S. and Edward M. Hubbard. 2001. "Synaesthesia: A window into perception, thought and language." *Journal of Consciousness Studies* 8(12), 3–34. https://philpapers.org/rec/RAMSA-5.

Ramsay, Timothy O., Richard T. Burnett, and Daniel Krewski. 2003. "The effect of concurvity in generalized additive models linking mortality to ambient particulate matter." *Epidemiology* 14(1), 18–23. doi: 10.1097/00001648-200301000-00009.

Sapir, Edward. 1929. "A study in phonetic symbolism." *Journal of Experimental Psychology* 12(3), 225–39. doi: 10.1037/h0070931.

Saussure, Ferdinand De. 1916. *Course in general linguistics*, edited by Charles Bally and Albert Sechehaye. Paris: Payot.

Schriefers, Herbert, Angela D. Friederici, and Uwe Rose. 1998. "Context effects in visual word recognition: Lexical relatedness and syntactic context." *Memory and Cognition* 26(6), 1292–1303. doi: 10.3758/BF03201201.

Shinohara, Kazuko and Shigeto Kawahara. 2010. "A cross-linguistic study of sound symbolism: The images of size." *Annual Meeting of the Berkeley Linguistics Society* 36(1), 396–410. doi: 10.3765/BLS.V36I1.3926.

Thompson, Patrick D and Zachary Estes. 2011. "Sound symbolic naming of novel objects is a graded function." *Quarterly journal of experimental psychology* 64(12), 2392–404. doi: 10.1080/17470218.2011.605898.

Tomaschek, Fabian, Peter Hendrix, and R. Harald Baayen. 2018. "Strategies for addressing collinearity in multivariate linguistic data." *Journal of Phonetics* 71, 249–67. doi: 10.1016/j.wocn.2018.09.004.

Uno, Ryoko, Kazuko Shinohara, Yuta Hosokawa, Naho Atsumi, Gakuji Kumagai, and Shigeto Kawahara. 2020. "What's in a villain's name?: Sound symbolic values of voiced obstruents and bilabial consonants." *Review of Cognitive Linguistics* 18(2), 428–57. doi: 10.1075/RCL.00066.UNO.

Vijver, Ruben van de and Dinah Baer-Henney. 2014. "Developing biases." *Frontiers in Psychology* 5, 1–8. doi: 10.3389/fpsyg.2014.00634.

Vitevitch, Michael S. and Paul A. Luce. 1998. "When words compete: Levels of processing in perception of spoken words." *Psychological Science* 9(4), 325–9. doi: 10.1111/1467-9280.00064.

Westbury, Chris, Geoff Hollis, David M. Sidhu, and Penny M. Pexman. 2018. "Weighing up the evidence for sound symbolism: Distributional properties predict cue strength." *Journal of Memory and Language* 99, 122–50. doi: 10.1016/J.JML.2017.09.006.

Wickham, Hadley. 2016. *ggplot2: Elegant graphics for data analysis*. New York: Springer.

Wood, Simon N. 2017. *Generalized additive models: An introduction with R*. Boca Raton: CRC Press. doi: 10.1201/9781315370279.