Sound symbolism in sighted and blind. The role of vision and orthography in sound-shape correspondences

Roberto Bottini[a,b,*], Marco Barilari[c], Olivier Collignon[a,c]

[a] *Center for Mind/Brain Sciences (CIMeC), University of Trento, Italy*

[b] *International School for Advanced Studies (SISSA), Trieste, Italy*

[c] *Institute of Psychology (IPSY) and Institute of Neuroscience (IONS), University of Louvain, Belgium*

ABSTRACT: Non-arbitrary sound-shape correspondences (SSC), such as the "bouba-kiki" effect, have been consistently ob- served across languages and together with other sound-symbolic phenomena challenge the classic linguistic dictum of the arbitrariness of the sign. Yet, it is unclear what makes a sound "round" or "spiky" to the human mind. Here we tested the hypothesis that visual experience is necessary for the emergence of SSC, supported by empirical evidence showing reduced SSC in visually impaired people. Results of two experiments comparing early blind and sighted individuals showed that SSC emerged strongly in both groups. Experiment 2, however, showed a partially different pattern of SSC in sighted and blind, that was mostly explained by a different effect of orthographic letter shape: The shape of written letters (spontaneously activated by spoken words) influenced SSC in the sighted, but not in the blind, who are exposed to an orthography (Braille) in which letters do not have spiky or round outlines. In sum, early blindness does not prevent the emergence of SSC, and differences between sighted and visually impaired people may be due the indirect influence (or lack thereof) of orthographic letter shape.

1. Introduction

Sound symbolism is the idea that some vocal sounds or phonemes carry meaning in non-arbitrary ways. A popular example is the shape-sound correspondence, according to which particular forms of nonsense words are consistently associated to particular unfamiliar shapes. In a seminal experiment, Wolfgang Khöler (Kohler, 1947) showed that people consistently matched the nonword "takete" to an image of a spiky object and the nonword "maluma" to an image of a rounded object (Fig. 1). More recently, this paradigm was popularized by Ramachandran and Hubbard (Ramachandran & Hubbard, 2001) using the nonwords *Bouba* and *Kiki*. Since then, the *bouba-kiki effect* has been replicated in many experiments (Ahlner & Zlatev, 2010; Aveyard, 2012; Maurer, Pathman, & Mondloch, 2006; Nielsen & Rendall, 2011; Westbury, 2005) across different cultures (Bremner et al., 2013; Chen, Huang, Woods, & Spence, 2016).

Despite extensive investigation, it is still unclear what drives the existence of sound-shape correspondences. This question is of particular interest since sound-symbolic associations (e.g. onomatopoeia, phonestemes, iconicity) are quite frequent across human languages (Blasi, Wichmann, Hammarström, Stadler, & Christiansen, 2016) and defies the classic dictum of the arbitrariness of the sign (de Saussure, 1959). Moreover, Sound symbolism facilitates word learning in both children (Imai et al., 2015; Imai, Kita, Nagumo, & Okada, 2008; Kantartzis, Imai, & Kita, 2011) and adults (Lockwood, Dingemanse, & Hagoort, 2016; Monaghan, Mattock, & Walker, 2012; Nygaard, Cook, & Namy, 2009) and may shed light on the processes of language evolution, where non-arbitrary relationships between sounds and objects may have served as a bootstrap for a more complex and largely arbitrary reference-system (Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015; Imai & Kita, 2014; Perlman, Dale, & Lupyan, 2015; Ramachandran & Hubbard, 2001).

What makes a sound "spiky" or "round"? Previous work (Nielsen & Rendall, 2013, 2011) has shown that sound-shape associations are carried both by consonants and vowels, with sound-symbolic patterns being relatively consistent across experiments (Sidhu & Pexman, 2017; Spence, 2011). For instance, sonorant consonants (e.g., /n/ /m/ /l/; e.g., as in "maluma") are consistently associated with continuous and round shapes. On the other hand, unvoiced stop consonants (e.g., /t/ /k/; e.g., as in "takete"), are consistently associated to irregular and spiky shapes. Indeed, the association sonorants-round and stops-spiky emerged in several experiments: from the classic maluma-takete effects

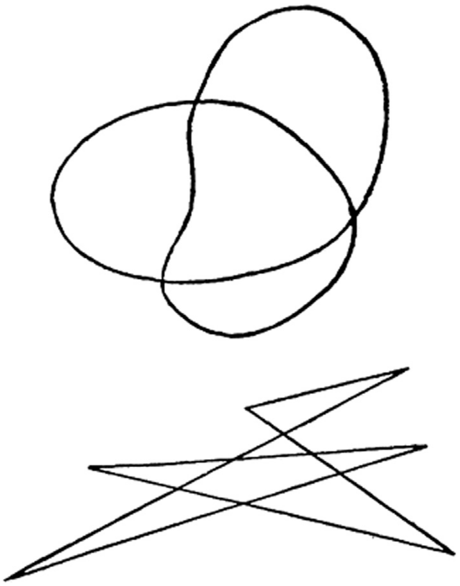

Fig. 1. Maluma and Takete in Köhler's original drawings.

with overt forced-choice paradigms (Kohler, 1947) to experiments testing implicit associations (Hung, Styles, & Hsieh, 2017; Parise & Spence, 2012; Westbury, 2005) or using several items and varying phonemes and visual outlines in a highly controlled fashion (Fort, Martin, & Peperkamp, 2015; McCormick, Kim, List, & Nygaard, 2015; Nielsen & Rendall, 2013; Westbury, Hollis, Sidhu, & Pexman, 2017). Other consonants, such as fricatives or affricates (/v/ /z/ /d͡ʒ/, /f/, /s/, /t͡ʃ/), seems to fall somewhere in the middle of the spiky-round continuum (Fort, Martin, & Peperkamp, 2015; McCormick et al., 2015), although their association pattern is more variable across experiments (e.g. Fort et al., 2015; Aveyard, 2012). A similar situation concerns voiced stop consonants (e.g., /b/, /d/, /g/), that show a mixed pattern of association with shape. Whereas the phoneme /b/ (as in "bouba") is normally highly associated with round shapes (Hung et al., 2017; Ramachandran & Hubbard, 2001; Westbury et al., 2017), the phoneme /d/ and /g/ do not show a clear pattern of association (Westbury et al., 2017). In the case of vowels, studies have consistently shown that back-vowels such as /o/ and /u/, are associated with round shapes, whereas front-vowels like /i/ and /e/, are associated with spiky shapes (Nielsen & Rendall, 2011; Westbury et al., 2017).

However, it is till unclear which features of the consonant and vowel phonemes (or their interaction) lead to the association with spiky or curvy shapes (Sidhu & Pexman, 2017; Spence, 2011). One possibility is that sound-shape associations are mediated by articulatory movements of the mouth and tongue during speaking. Phonemes are mapped onto the motor processes needed to produce them, and these dynamic movements carry similarities with certain shapes (Ramachandran & Hubbard, 2001). For instance, the sharp inflection of the tongue on the palate when producing the consonant /t/, together with the abrupt obstruction of airflow (as in "takete") may carry similarities with spiky irregular outlines (Westbury, 2005). Whereas the round-mouth articulation of the /oo/ vowel sound (as in "bouba") can be iconic to rounded shapes (Ramachandran & Hubbard, 2001). It is still unclear, however, whether these maps between sound, articulation and shape would be learned during development, due to the regular coupling between articulatory and phonemic expressions in speech (Deroy & Spence, 2013), or are largely innate (Ramachandran & Hubbard, 2001), possibly justified by an initial state of "neonatal synesthesia" (Maurer & Mondloch, 2005) according to which the sensory experience of infants is less diversified, and sensory stimulations triggers blended sensations similar to the one reported by adult synesthetes (Maurer et al., 2006).

Another related proposition focus on the relationship between acoustic features of vocal sounds and shapes, independently of manner of articulation (Nielsen & Rendall, 2013; Nielsen & Rendall, 2011). For instance, spiky shapes and smooth shapes may carry some intrinsic similarity with the frequency (e.g., pitch) and dynamic pattern of some words and phonemes (Knoeferle, Li, Maggioni, & Spence, 2017; Nielsen & Rendall, 2013; Nielsen & Rendall, 2011). High pitch sounds (as in the case of front vowels or unvoiced consonants; Kirby & Ladd, 2016) may share spectral properties with jagged (high-frequency) visual shapes (Chen et al., 2016; Shang & Styles, 2017), and voiceless stops consonants, characterized by complete silence followed by an abrupt burst of sound, may be iconic to the abrupt directional changes of irregular outlines (Ramachandran & Hubbard, 2001). As in the case of articulatory patterns, the associations between acoustic features and shape could be innate, possibly based on synesthetic processes (Maurer & Mondloch, 2005) or motivated by evolutionary processes shared with other animals (Ohala, 1997; Sidhu & Pexman, 2017).

Other explanations of sound-shape correspondences focus instead on indexical associations based on environmental co-occurrences. For this account, there is nothing intrinsically similar (i.e., iconic) between sounds and shapes: They just happen to co-occur regularly in our experience. For instance, pointy objects may produce higher-pitch and strident sounds (e.g., when they fall or we interact with them), whereas rounded objects produce more mellifluous sounds with less abrupt transitions (Fryer, Freeman, & Pring, 2014; McCormick et al., 2015).

Despite the large body of experimental evidence about sound-shape associations and sound-symbolism in general, it is currently impossible to definitely choose one of the proposed mechanisms over the other (Sidhu & Pexman, 2017; Spence, 2011). This state of affairs is largely due to the fact that the majority of experiments have not been designed to compare alternative theories (Sidhu & Pexman, 2017).

One possibility to shed some light on the processes underlying sound-shape associations could be comparing people that have a drastically different experience of the world, and see how this affect sound-symbolic correspondences. A particularly good model system for this aim is visual deprivation. Blind individuals lack vision, which is the primary sense for shape perception in sighted people (Denys et al., 2004), and one of the primary sensory modalities to collect statistics from the surrounding environment (Deroy, Fasiello, Hayward, Auvray, et al., 2016; Fryer et al., 2014; Smith, Colunga, & Yoshida, 2010).

Recent findings have boosted the interest on blindness as a model system to study sound symbolism. Fryer and colleagues (Fryer et al., 2014) tested sighted people and a heterogeneous group of visually impaired subjects (including early blind, late blind and partially sighted) in a haptic version of the bouba-kiki task. Although a sound-symbolic effect emerged in both populations, the effect was significantly lower in the visually impaired group. Interestingly, the subgroup of 6 early blind did perform at chance in the task (Fryer et al., 2014). This result has been followed up by experiments testing cross-modal correspondences in larger groups of Early and Late Blind individuals (Barilari, de Heering, Crollen, Collignon, & Bottini, 2018; Deroy, Fasiello, Hayward, & Auvray, 2016). One study (Deroy, Fasiello, Hayward, Auvray, et al., 2016) explored the relationship between pitch and tactile height, using an implicit association task. Whereas sighted showed the classic association between higher pitch and higher spatial locations (and vice versa), Early and Late Blind showed a reduced and non significant pattern of cross-modal correspondence, suggesting that pitch-height associations are vision-dependent. These data suggests that lack of vision may have an impact on the development of at least some form of sound symbolism and, more generally, cross-modal associations. Of particular interest, the hypothesized absence of the bouba-kiki effect in early blind suggests that early visual experience may be necessary for this association to emerge (Fryer et al., 2014). For instance, it is possible that intrinsic similarities between sounds and shapes, that may lead to the bouba-kiki effect (Nielsen & Rendall, 2011; Ramachandran & Hubbard, 2001), cannot be grasped via a uniquely haptic experience of shapes. Alternatively, critical sound-shape associations in the environment (e.g., lips articulation during speech; Ramachandran & Hubbard, 2001) may be specific of visual experience (Fryer et al., 2014). In any case, evidence that vision plays a pivotal role in the development of sound-shape associations would falsify strong innatist accounts related to the hypothesis of neonatal synesthesia (Maurer & Mondloch, 2005) and help to pinpoint the (visual) features of shape processing and/or environmental contingencies that may be fundamental for the development of sound symbolic correspondences.

2. Experiment 1

In Experiment 1 we proposed to sighted and early blind participants an haptic version of the maluma-takete task originally designed by Kohler (Kohler, 1947). Our task was almost identical to (and inspired by) the task performed by Fryer and colleagues in their seminal study (Fryer et al., 2014), with the main exceptions that we used the non-words *maluma* and *takete* (instead of *bouba* and *kiki*) and we tested a larger sample of early blind people (n = 30).

2.1. Methods

2.1.1. Participants

Sixty participants completed the experiment in exchange for payment: 30 early blind (EB) and 30 sighted controls (SC). Participants in the EB and SC group were matched pairwise for gender (17 females), age (EB = 36.27, sd = 10.86; SC = 36.70, sd = 11.62) and years of education (EB = 14.10, sd = 3.09; SC = 14.43, sd = 2.92; see also Supplementary materials, Table 1). All participants were Italian native speakers and were blindfolded during the task. Participants in the EB group lost completely their sight at birth or before 4 years of age and were totally blind or had only rudimentary sensitivity for brightness differences. All of them could not have visually perceived shapes and/or have had visual memory of shapes. In all cases, blindness was attributed to peripheral deficits with no additional neurological problems. All our blind participants had no experience of visual reading and were fluent in Braille (see Supplementary material, S7, for further information about how we acquired these information).

The sample size was decided based on the amount of early blind participants we could reach and the strength of the effect. Although the experiment had only 4 trials, the bouba-kiki effect is known to be fairly strong, with ~80–90% of participants choosing the typical shape-sound associations (Styles & Gawne, 2017), suggesting that it should be highly replicable, even with small samples. Moreover, we substantially increased the number of early blind participants (from 6 to 30) compared to the experiment that we intended to conceptually replicate (Fryer et al., 2014). The ethical committee of the University of Trento approved this study and all participants were naive with respect to the purpose of the experiment.

2.1.2. Material

Stimuli were four pairs of shapes modeled after Fryer et al. (2014). However, whereas Fryer et al. stimuli were made of different materials, all our stimuli were custom made in solid resin with a 3-D printer (Fig. 2). Pair A and pair B were designed to mimic Köhler line drawings of *maluma* and *takete*. Objects in Pair A were 3D shapes, one smooth and bulb-shaped and the other spiky and irregular in all dimensions. Pair B consisted of 2D shapes, one rounded and one spiky all over. In Pair C there were two discs identical in shape but different in texture (smooth Vs. cross-hatched) and in pair D we had two spheres, one smooth and the other spiky all over. Each pair of objects was contained in a box measuring 300 mm × 17 mm × 12 mm.

2.1.3. Procedure

Instructions were provided to the subjects orally, according to a predefined script (see Supplementary material; S4). The four boxes were presented on the table in front of the person, one at the time. Participants, who were blindfolded before entering the room, were asked to explore the two objects inside the box with both hands, and to bring out either *maluma* or *takete*. The experiment consisted of four trials, one for each object pair. The order of trials was determined according to four different lists (ABCD, BCDA, CDAB, DABC) counterbalanced across subjects. In half of the trials participants were asked to bring out "*maluma*" and in the other half to bring out "*takete*" (alternated across trials). The counterbalancing led to a total of 8 (4 × 2) different rotations. For each trial, subjects scored 1 point if they mapped word and shape in the expected way (e.g. spiky object as *takete*), otherwise they scored zero. The total score, thus, ranged from 0 to 4. All participants were debriefed after the experiment. The interview was conducted orally and we asked: "How did you choose which one was *maluma* and which one was *takete*?" (see supplementary material; S1).

2.2. Results

Both groups showed a sound symbolic effect mapping shapes to words in the expected manner (Fig. 3). The SC group scored on average 3.63/4. Twenty-five out of thirty (83.33%) chose as expected for all 4 pairs, selecting rounded objects as *maluma* and spiky objects as *takete*; one subject choose consistently in the opposite direction and the remaining 4/30 (13.33%) were inconsistent, choosing a rounded shape sometimes as *takete* and sometimes as *maluma*. The EB group had an average score of 3.13/4. Twenty-two participants out of thirty (73.33%) choose as expected in all the 4 trials, 5/30 (16.67%) chose consistently in the opposite way and the 3/30 (10%) were inconsistent.

A one sample *t*-test showed that the score of each group was significantly higher than chance, for the SC group: *t* (29) = 9.227, p = < 0.001, Cohen's d = 1.69, and for EB group: *t* (29) = 3.954, p = < 0.001, Cohen's d = 0.72. A two-samples *t*-test showed no differences across groups (*t*-test: t(58) = 1.486, p = 0.143, Cohen's d = 0.39). To further qualify this null results we run a Bayesian *t*-test (JASP Team, 2017) which revealed very little (anecdotal) evidence against the null hypothesis (BF10 = 0.659; Jeffreys, 1961; Kass & Raftery, 1995). However, since a Shapiro-Wilk test signaled a deviation from normality ($W > 0.44$, $p < .001$) we confirmed this null result with a non-parametric permutation test resampling our data 1000 times. Also in this case the difference between groups was not significant (p = .17, Confidence interval = [0.14, 0.18]).

2.3. Discussion

Early loss of vision does not prevent the development of shape-sound correspondences. Early Blind individuals showed a significant bouba-kiki effect that was statistically indistinguishable from their sighted counterpart. This result, fails to provide strong evidence against the universality and innateness of sound-symbolic associations, and suggests that the iconic and/or correlational ground on which shape-sound associations are constructed is relatively resilient to early visual deprivation. In other words, if the bouba-kiki effect is mostly driven by intrinsic similarities between sounds and shapes (Nielsen & Rendall, 2011; Ramachandran & Hubbard, 2001), these similarities can be equally established without vision (i.e., via haptic experience of shape). On the other hand, if sound-shape correspondences are mostly driven by statistical associations in the environment (Fryer et al., 2014), these associations are not precluded (at least not completely) by visual deprivation.

However, if our results exclude the hypothesis that blindness prevents sound-shape correspondences, they do not exclude the possibility that it can influence them in interesting ways. Indeed, as the result of Fryer et al. (2014) suggest, visual impairment may in some cases modulate sound-symbolic patterns. In our next experiment we aimed to test a hypothesis that predicts an impact of early blindness on shape-sound associations and can account, at least in part, for the

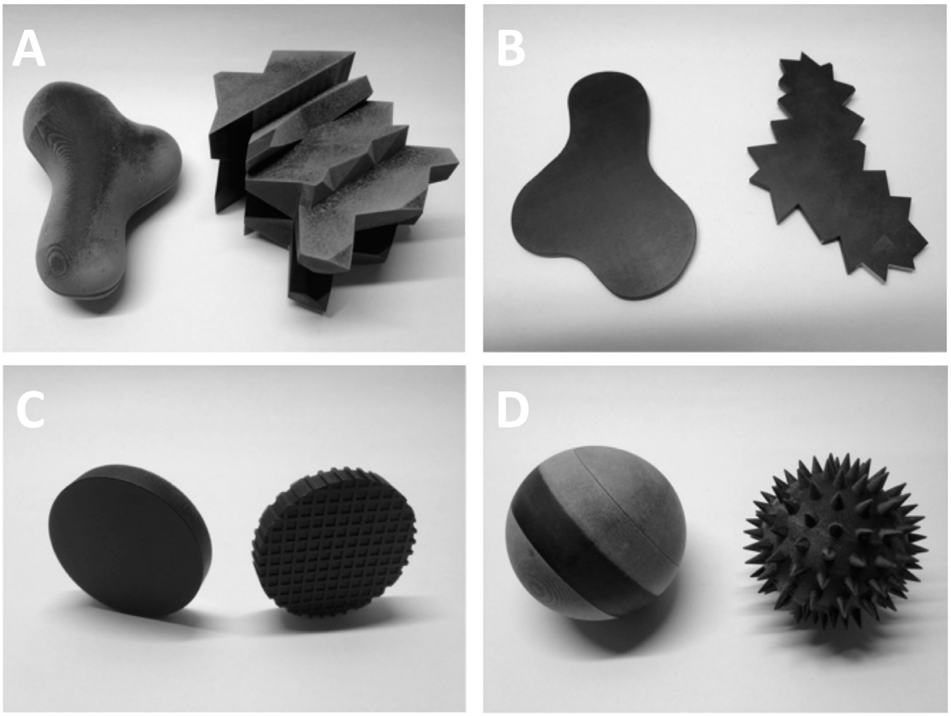

Fig. 2. Object Pairs used in Experiment 1.

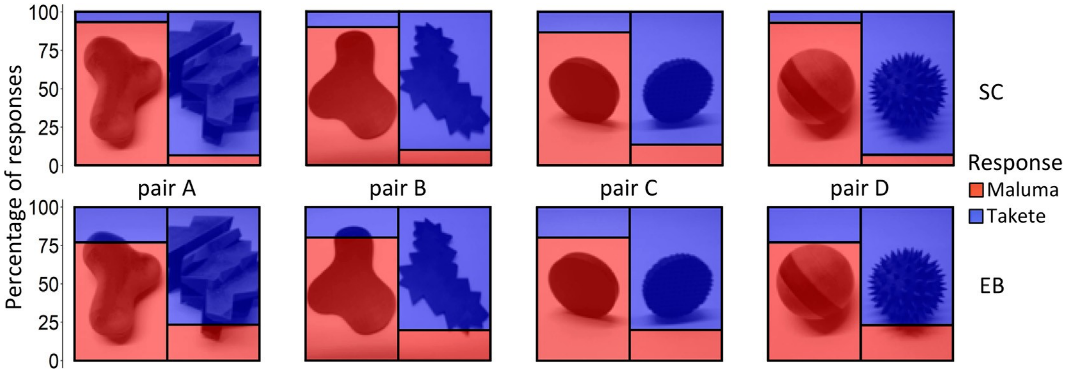

Fig. 3. Percentage of maluma and takete responses for each object in the sighted (first row) and early blind (second row) group.

discrepancies between our experiment and Fryer's one.

The hypothesis is that shape-sound correspondences are influenced by the shape of written letters. Consonants and vowels are not only associated with a sound, but, at least in the mind of literate people, also with an orthographic sign. The role of orthography in sound symbolism has been mostly approached as a source of confounds to be ruled out (Hung et al., 2017; Sidhu, Pexman, & Saint-Aubin, 2016; Westbury, 2005), showing that shape-sound associations *can* emerge independently of orthography. Studies on pre-literate children (Maurer et al., 2006), babies (Ozturk, Krehm, & Vouloumanos, 2012) and adults from populations without a writing system (Bremner et al., 2013) successfully proved this point.

However, there is also evidence that letter shape can modulate sound-symbolic judgments (Doyle & Bottomley, 2011; see also Turoman & Styles, 2017, suggesting how, historically, the shape of written letters may have been influenced by sound-symbolic mechanisms). For instance, in one of the few systematic studies of orthographic influence on sound symbolism, Cuskley and colleagues (Cuskley, Simner, & Kirby, 2015) showed that orthographic spykeness (i.e. the angularity of letters shape) was actually the strongest predictor of the bouba-kiki effect when compared to phonological factors such as voicing and manner of articulation, even with aurally presented words. Nevertheless, the relative contribution of orthographic and phonological factors cannot be clearly drawn from Cuskley et al. (2015) given that: (i) only 8 items were used in their experiment; (ii) voicing and manner of articulation were not orthogonalized with respect to orthography; and (iii) the articulatory/acoustic spectrum was not fully covered (e.g. there were no sonorant ('round') consonants). Moreover, the hypothesis that orthography is the main driver behind sound-shape correspondences can hardly explain the results with toddlers, babies and adults that were never exposed to an orthographic system (Bremner et al., 2013; Maurer et al., 2006; Ozturk et al., 2012).

Here, instead, we suggest that (i) Sound-symbolic relationships between sounds and shape are sufficient for the emergence of the typical bouba-kiki effect; (ii) However, when listening to words literate people spontaneously activate the orthographic representation of these words (Chéreau, Gaskell, & Dumay, 2007; Ziegler & Ferrand, 1998), and the shape of written letters influences cross-modal mappings between shape and sound. A model of sound-symbolic correspondences that includes orthographic shape may explain the discrepancies between our Experiment 1 and Fryer et al. results (2014). Indeed, in their tactile experiment, Fryer et al. used the nonwords "bouba" and "kiki", which are

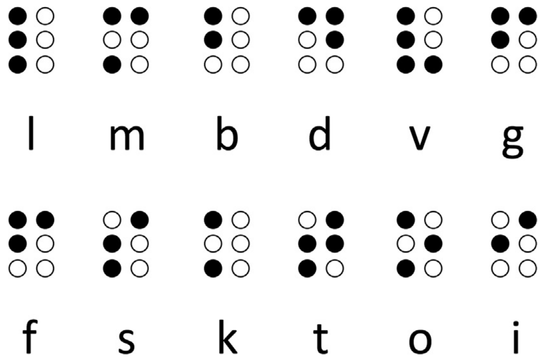

Fig. 4. Comparison between letters in Braille and Latin alphabet.

both composed of stop consonants (/b/ and /k/), whose difference in terms of spikiness may be related in part to their orthographic shape. Such orthographic influence, may be absent in visually impaired people, especially if they are blind from early in life. Although early blind people speak, read and write using the same language than their sighted counterparts (e.g., Italian, or English), they are exposed to a very different orthography provided by the tactile alphabet Braille. Contrary to written letters, Braille letters are composed by dots arranged in a 3 × 2 grid, without spiky or round edges as letters in the written alphabet (see Fig. 4). Since braille alphabet cannot provide letter representations in terms of roundness or spikiness, the pattern of sound-symbolic association shown by blind and sighted people may be relatively different.

3. Experiment 2

In Experiment 2 we tested whether sighted and early blind people show a different pattern of sound symbolic correspondences and whether, such a difference can be explained (at least in part) by different orthographic experience. In this experiment, early blind people and sighted controls listened to several pseudowords and decided whether each word sounded “round” or “spiky”.

3.1. Methods

3.1.1. Participants

Thirty-six participants completed the experiment in exchange of payment: 18 people with early blindness (EB) and 18 sighted controls (SC). Participants in the EB and SC group were matched pairwise for gender (9 females), age (EB = 38.56, sd = 13.07; SC = 39.56, sd = 12.91) and years of education (EB = 15.22, sd = 3.26; SC = 15.83, sd = 2.96; see also Supplementary material, Table 1). All participants were Italian native speakers. Half of the participants (9 EB and 9 SC) participated previously to Experiment 1 (the distance between Exp.1 and Exp. 2 was of at least 4 months). Participants in the EB group were totally blind or had only rudimentary sensitivity for brightness differences. In all cases, blindness was attributed to peripheral deficits with no additional neurological problems. All our blind participants had no experience of visual reading and were fluent in braille.

Given the lack of previous estimation of the effect size of orthographical factors on sound symbolism (especially when the articulatory/acoustic factors are taken into account), we did not have a precise estimate on which we could base our sample size. However, the previous systematic study reporting orthographic influence on phoneme-shape correspondences tested 41 participants, 4 trial per participants (Cuskley et al., 2015). Given the relatively large number of trials (n = 240) in our experiment, we thought that a standard sample size of 18–20 participants could be acceptable. As it is often the case for research with special populations, our sample size choice was constrained by the amount of early blind participants that we could reach and that would agree to complete a relatively long experiment (~100 min). The ethical committee of the University of Trento approved this study and all participants were naive with respect to the purpose of the experiment.

3.1.2. Material

Stimuli consisted in 240 pseudo-words audio recorded by a male Italian native speaker and edited into separated audio files with the same auditory properties (44,100 Hz, 32 bit, mono, 78 dB of intensity). We systematically sampled the acoustic and phonemic space using pseudo-words that were composed by one of 5 classes of consonants (see McCormick et al., 2015 for a similar design): (1) Sonorant; (2) Voiced Af/Fricative; (3) Unvoiced Af/Fricative; (4) Voiced Stop; (5) Unvoiced Stop.

All words have a CVCVCV structure based on the two pseudo-words used in Exp.1 (see see Supplementary material, Table3, for the complete list). Each stimulus contained consonants from one of five different classes: Voiced Sonorants (/n/, /m/, /l/), Voiced Af/Fricative (/v/, /z/, /d͡ʒ/), Unvoiced Af/Fricative (/f/, /s/, /t͡ʃ/), Voiced Stop (/d/, /g/, /b/) or Unvoiced Stop (/t/, /k/, /p/). There were 48 words for each consonant class, meaning that all the consonants in a given word came from the same class. All words had equal consonants in the first and third position, and a different consonant in the second position (as in *maluma* and *takete*). Each consonant appeared an equal number of times across the stimuli.

For each consonant class, half of the stimuli contained back/rounded vowels (/o/, /u/), and the other half contained front/unrounded vowels (/e/, /i/). In half of the stimuli the same vowel appeared in the first and third position (e.g. mulomu), and for the other half in the second and third position (e.g. tekiti).

3.1.3. Procedure

All the participants completed the survey on-line using the platform SurveyGizmo™. The survey was made accessible for visually impaired people and its accessibility was validated by two blind participants (not included in the sample) before starting data collection. Participants were strongly encouraged to listen to the stimuli through headphones for a better sound quality. Each stimulus in the set was presented aurally one by one and participants were asked to focus only on the sound of the pseudo-word and to classify it as round or pointy. Stimuli duration mean was 0.71 s (sd = 0.06 sec) plus 2.5 s of silence before stimulus onset to let the participants focus on the task after having pressed the play button. Participants had the possibility to listen to the stimuli more than once without any time constriction and the trial presentation was self-paced. To choose the response, participants were instructed to open a window menu and select only one of the two options, round or spiky. It follows that the participants of the blind group used a voice synthesizer to navigate the web page were the survey was administered. The order of the stimuli was pseudo-randomized according to two different lists counterbalanced across subjects. The position of the two responses (i.e. round – pointy or pointy – round) was also counterbalanced across subjects. The counterbalancing led to a total of 4 (2 × 2) different versions of the survey.

3.1.4. Analyses and predictors

Analyses were conducted using a Mixed Effect GLM with binomial distribution. Data were analyzed using R (R Core Team, 2013) and the package lme4 (Bates, Mächler, Bolker, & Walker, 2014). Random effects were the intercepts and slopes for subjects and items, whereas fixed effects were Consonant Type (3 levels, Sonorants, Fricative/Affricates, Stops), Voicing (Voiced, Unvoiced), Vowel Type (round, unround), Orthographic Spikiness, and Group (EB, SC). Subject-wise and item-wise slopes, in the random effect structures, corresponded to the highest-order combination of within-subject factors subsumed by each interaction of interest (Barr, 2013).

The regressor Orthographic Spikiness (OS from here) was created as follows. Two separate groups (n = 20 per group) completed an on-line survey in which they had to rate the visual spikiness of single letters from 1 to 7 (1 = not spiky, 7 = very spiky). Since we had no clear a-priori reason to select a particular letter notation, one group saw the letters in their upper case (e.g. A, B, C) and the other group in lower case (e.g. a, b, c) (See Supplementary materials; S5). We then averaged the ratings to obtain the mean spikiness rating for each letter.

In a different survey, 10 additional sighted participants listened to some of the pseudowords used in the main experiments. These were pseudowords that according to Italian orthography could have been written in more than one way (e.g. /takete/ can be *takete* or *tachete*). Participants were asked to listen each pseudoword and then write it down, in order to confirm empirically the orthographical alternatives (See Supplementary materials; S6).

Based on the rating of OS for each letter we then computed the average orthographic spikiness of each pseudoword in our experiment. When a pseudoword had more than one orthographic alternative, we computed its final OS by averaging the OS of all the possible alternatives (as indicated by the survey). Values of OS were standardized before running statistics to allow a meaningful comparison between odd ratios.

Importantly, our measure of OS did not include the shape of vowels. That is because vowel's OS is highly correlated with their articulatory/auditory spikiness: Back vowels graphemes ("o", "u") are more round than front vowels ones ("e", "i"). In our attempt to make orthographic and articulatory/auditory spikiness as much orthogonal as possible, including vowels in the count of OS seemed counterproductive and possibly misleading.

3.2. Results

We tested for the effect of acoustic and orthographic factors by entering "Spiky" and "Round" responses in a Mixed Effect Generalized Linear Model with binomial distribution. Random effects were the intercept and slopes of subjects and items (Barr, 2013; Barr, Levy, Scheepers, & Tily, 2013). Fixed effects were Consonant Type (3 levels, ordered for consonant class as Sonorant, Af/Fricative and Stops), Voicing (voiced, unvoiced), Vowel Type (back, front), Orthographic Spikiness, and Group (EB, SC). The model tested for the main effects of each factor and their interaction with the factor Group. Odd-ratio and Upper/Lower levels of the confidence interval are reported as a measure of effect size.

Since this model failed to converge, we ran the same maximal model with random slopes and intercepts for subjects only (Mathôt, Grainger, and Strijkers, 2017) and the model finally converged. However, the pattern of significance was the same of the non-converging model.

We found, as expected, a large effect of Consonant Type ($z = 6.29$, $p < .001$, Odd-ratio = 15.20, LL = 6.51, UL = 35.47), Voicing ($z = -3.60$, $p < .001$, Odd-ratio = 0.47, LL = 0.31, UL = 0.71) and Vowel Type ($z = 3.60$, $p < .001$, Odd-ratio = 4.72, LI = 2.03, HI = 10.99).

As shown in Fig. 5, sonorant consonants were associated with round responses, stop consonants with spiky responses, and fricative/affricatives were somewhat in the middle. Similarly, as shown in Fig. 6, back vowels (/o/, /u/) were associated with round responses and front (/e/, /i/) vowels with spiky responses. The main effect of Orthographic spikiness ($z = -0.80$, $p = .42$, Odd-ratio = 0.94, LL = 0.81, UL = 1.09) and Group ($z = -0.55$, $p = .58$, Odd-ratio = 0.76, LL = 0.29, UL = 1.99) did not reach significance.

However, we found a significant interaction between Consonant Type and Group ($z = -2.31$, $p = .02$, Odd-ratio = 0.34, LL = 0.14, UL = 0.85), and, most importantly, between Orthographic Spikiness and Group ($z = 3.24$, $p = .001$, Odd-ratio = 1.42, LL = 1.15 UL = 1.75).

To better articulate this latter interaction we divided the database in sighted and blind group and we ran two separate models, one per each group with Consonant Type, Voicing, Vowel Type and Orthographic Spikiness as regressors, and intercept and slope of subjects as random effects. In the EB group the factor Orthographic Spikiness did not reach significance ($z = -0.92$, $p = .36$, Odd ratio = 0.94, LL = 0.82 UL = 1.07).

In contrast, in the SC group the factor Orthographic Spikiness was highly significant ($z = 3.48$, $p < .001$, Odd ratio = 1.33, LL = 1.13 UL = 1.55), suggesting that letter shape was an important factor in determining the judgments of sighted people, and qualifying the interaction between groups.

3.3. Discussion

As predicted, the orthographic shape of written letters predicted sound-shape correspondences in sighted but not in early blind people. This finding suggests that the reduced bouba-kiki effect previously reported in visually impaired people (Fryer et al., 2014) may be due, at least in part, to the role that orthography plays in the emergence of this effect. Fig. 7 clearly exemplifies this by reporting difference in spikiness ratings between stop unvoiced (e.g. /k/ as in kiki) and stop voiced (e.g. /b/ as in bouba) consonants, for each subject. This difference is normally small in the EB group and larger in the SC group: this is in line both with Fryer et al. results and our effect of orthography.

However, both in sighted and blind people, the effect of orthography was clearly smaller compared to the effect of voicing and manner of articulation. This results is in contrast with previous findings suggesting a prominence of orthographic factors (Cuskley et al., 2015) and in line with results showing bouba-kiki effects with babies (Ozturk et al., 2012) and adults (Bremner et al., 2013) that were never exposed to orthography.

4. General discussion

Sound-symbolic association between shape and phonemes is a robust cross-cultural phenomenon that defies the arbitrariness of the sign traditionally endorsed by linguistic theories (de Saussure, 1959). The mechanisms of this cross-modal association remain debated, including the role of visual experience in shaping sound-symbolic patterns. Do the intrinsic similarities between sounds and shapes, that may lead to the bouba-kiki effect (Nielsen & Rendall, 2011; Ramachandran & Hubbard, 2001), can be grasped via haptic experience of shapes? Alternatively, if sound-shape correspondences are mostly driven by statistical associations in the environment (Fryer et al., 2014), are these associations precluded by visual deprivation? Preliminary data on a small group (n = 6) of early blind individuals who failed to show the classic bouba-kiki effect, suggested indeed that blindness may prevent sound-shape associations by acting on one of these mechanisms (Fryer et al., 2014). However, in our first experiment we showed that early blindness does not prevent the emergence of sound-shape correspondences. A large group of sighted (n = 30) and early blind people (n = 30) were engaged in a haptic version of a classic sound-shape correspondence task (Fryer et al., 2014; Kohler, 1947): Both sighted and blind associated the sound /maluma/ with round shapes and the sound /takete/ with spiky shapes, with no statistically significant difference between groups.

In experiment 2, we tested further the strength of sound-symbolic associations in sighted and early blind people, by asking participants to rate as 'spiky' or 'round' several nonwords carrying different acoustic, articulatory and orthographic features. This paradigm allowed us to test for more subtle differences between sound-symbolic patterns in the two populations, assessing the relative effects of acoustic and orthographic factors. In particular, we predicted that the shape of written letters could influence sound-shape associations in sighted but not in blind people, who are exposed to a different orthography (Braille) that does not have sharp angles or curvy outlines.

Results confirmed that both sighted and blind show a robust sound-

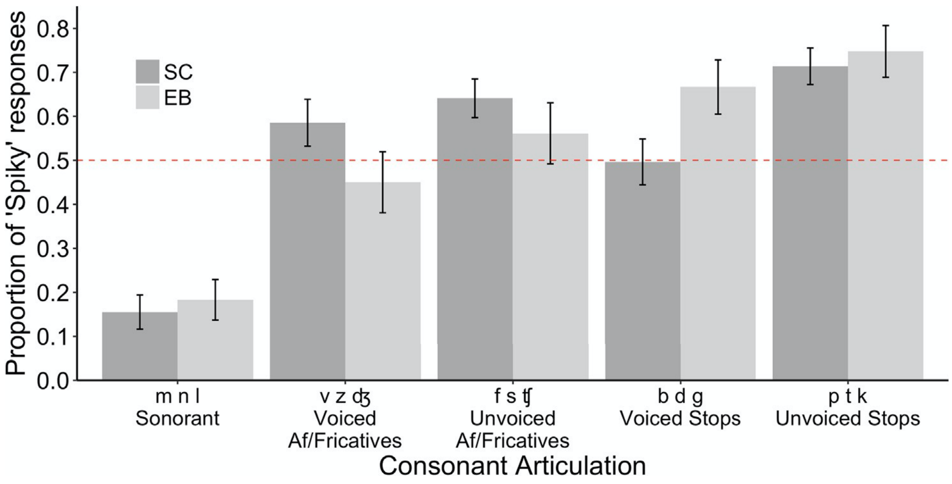

Fig. 5. Proportions of ‘spiky’ responses in Experiment 2, arranged by consonant class, for sighted (SC) and blind (EB). The response pattern provided by SC is very similar to the one reported in McCormick et al. (2015), whereas EB people show a partially different pattern. Dashed line is chance level. Error bars represent ± 1 SEM.

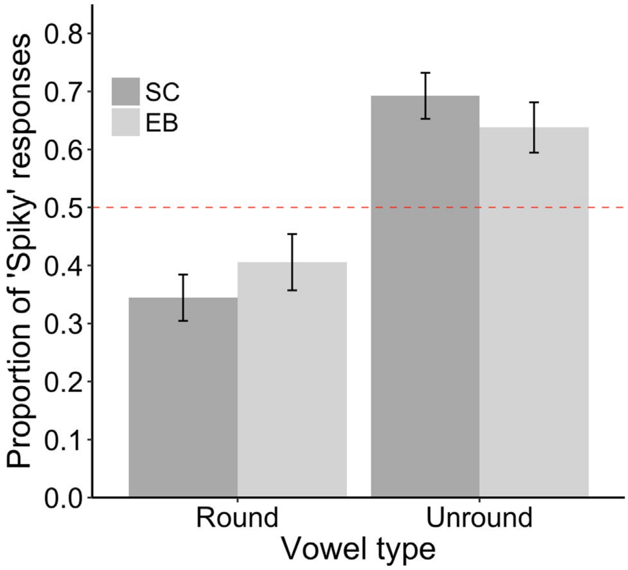

Fig. 6. Proportion of ‘spiky’ responses in Experiment 2 aggregated by vowel type. Dashed bar is chance level. Error bars represent ± 1 SEM.

symbolic association between phonemes and shape, largely in line with previous reports (Fort et al., 2015; McCormick et al., 2015; Nielsen & Rendall, 2011; Westbury et al., 2017). As predicted by models based on acoustic and articulatory features of phonemes (McCormick et al., 2015; Nielsen & Rendall, 2011; Westbury et al., 2017) both manner of articulations and voicing influenced sound-shape correspondences, in both groups: Sonorant consonants were considered rather round, stop consonants rather spiky, and frivatives/affricates somewhere in the middle (Fort et al., 2015; McCormick et al., 2015); Voiced consonants were generally considered more round than unvoiced consonants (D’Onofrio, 2014; McCormick et al., 2015). However, whereas the SC sound-symbolic pattern was influenced by the orthographic shape of letters, EB people showed no influence of orthography.

This state of affairs can reconcile our result with previous results showing a reduced bouba-kiki effect in visually impaired people compared to normally sighted. That is, /b/ as in bouba and /k/ as in kiki are both stop consonants, that may sound rather spiky in some cases, especially when compared with sonorant consonants (/l/, /m/, /n/; see for instance McCormick et al., 2015). The spiky outline of the letter ‘k’, and the round outline of letter ‘b’, however, can increase their difference in perceived spikiness in sighted compared to blind people who are mostly exposed to braille orthography (Fig. 7). This may explain a decreased (but still significant) bouba-kiki effect in visually impaired people compared to their fully-sighted counterpart, as previously reported (Fryer et al., 2014). Our results suggest that the primary factor differentiating sound-shape associations between sighted and blind may

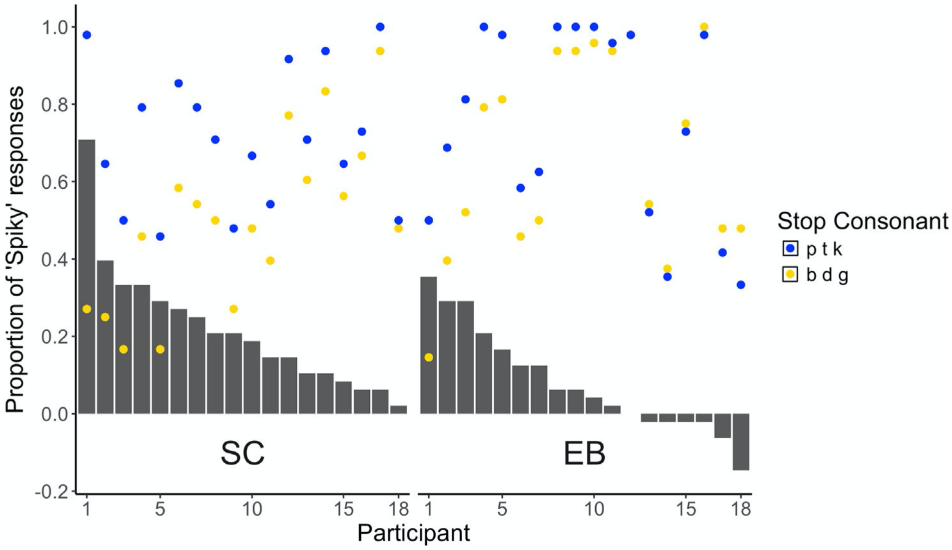

Fig. 7. Spikiness difference (bars) between unvoiced (blue dots) and voiced (yellow dots) stop consonants for each participant in the two groups. From left to right, the participants are ordered according to the strength of the effect. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

be orthographic shape.

Nevertheless, it is possible that differences in shape-sound correspondences between sighted and blind are also modulated by non-orthographic factors that are still not specified. Indeed, in our analysis, we found an interaction between Group (SC, EB) and Consonant type (sonorant, Af/Fricative, stop), above and beyond the interaction between Group and Orthographic Spikiness. One possible explanation is that our characterization of orthographic spikiness could not capture this phenomenon in its entirety, leaving out some subtle orthographic differences that otherwise would account for the portion of variance currently explained (spuriously) by consonant classes.

Alternatively, early blindness can impact directly on some specific sound-symbolic mechanisms that remain unspecified in our analysis. For instance, a recent paper published while our paper was under revision, showed that, contrary to sighted, early blind ($n = 32$) failed to associate high-pitch tones with jagged shapes and low-pitch tones with round shapes (Hamilton-Fletcher et al., 2018). Indeed, there is a large literature suggesting that low pitch is associated with larger/darker/lower/rounded characteristics of objects and events (Sidhu & Pexman, 2017; Spence, 2011), and to the extent that sound-shape correspondences such as the bouba-kiki effect are based on pitch-shape associations (Chandran, Banerjee, & Ghosh, 2017; Nielsen & Rendall, 2011; Sidhu & Pexman, 2017), they should be different between sighted and early blind. In our data, the modulatory effect of pitch (and its interaction with blindness) can be controlled in two ways, previously assessed in the literature (Chandran et al., 2017; McCormick et al., 2015). First, voiceless consonants usually induce higher pitch than voiced ones (Kirby & Ladd, 2016), from which their association with spiky and round shapes, respectively, may derive (Bar & Neta, 2006; Ohala, 1997; Sidhu & Pexman, 2017). However, the interaction between voicing and blindness was not significant in our model, providing no evidence for a different effect of voicing (and pitch) across groups.[1]

Second, one recent paper (Chandran et al., 2017) have suggested that the bouba-kiki effect may be due to a larger presence of high frequency components in "spiky" compared to "round" word. The relative power of high-frequency bands in a spoken word can be indexed by a High Frequency Fraction (HFF; Chandran et al., 2017) that can be used to predict associations with spiky or round shapes. However, when this predictor was entered as a regressor in our data (See Supplementary material; S3), we could not find an interaction between HFF and Group, whereas the interaction between Group and Orthographic Spykeness remained highly significant.

However, although we could not find evidence in our data that pitch-shape correspondences change as a function of visual deprivation, this is a hypothesis that deserves further exploration. Moreover, in the Hamilton-Fletcher experiments (Hamilton-Fletcher et al., 2018) participants judged the association between pitch and 3D shapes that they could touch (as in our Experiment 1), whereas in our Experiment 2 participants associated sounds to their own mental representation of "spikiness" and "curviness" that may be somehow different between sighted and blind. This is a difference between the two experiments that should be taken into account. On the other hand, it is also worth considering that Hamilton-Fletcher et al. (2018) did not debrief their participants about the strategies adopted during their task (which is similar to our Experiment 1). This may be an important detail, since both in our Experiment 1 (see Supplementary material; S1) and in Fryer et al. experiment (2014), blind individuals indulged more often (compared to sighted) in semantic associations instead of focusing on the formal (acoustic) properties of the words. For instance: "Kiki is a female name so I chose the rounded shape as Kiki", or "Takete reminded me the word hill ('tacco' in Italian), so I associated it to the round shape similar to a foot". The choice to rely on more semantic/conceptual strategies, or object-object relationships (Fryer et al., 2014) may be one of the causes of the lower phoneme-shape and pitch-shape associations in EB people, or a mere consequence of it. In this regard, future studies testing more implicit associations, with paradigms that leave little space to strategic approaches (Parise & Spence, 2012; Westbury, 2005), would be very helpful. In sum, further investigations are needed to better characterize the impact of blindness on sound symbolysm and cross-modal correspondences in general.

In conclusion, by demonstrating a solid bouba-kiki effect in early blind, our results open to the intriguing possibility that people with early blindness may be more sensitive to (some) iconic features of language compared to sighted. One of the major problem in language learning is referential ambiguity, meaning the difficulty of attributing uttered words to their referents in highly cluttered environments. One can imagine, for instance, the number of possible referents for a 1 y.o. language-learner when someone says the word "napkin" on a fully set table. Experimental evidence suggests that the ambiguity problem may be reduced, from the first year of life, by the frequency distribution of objects in infants' egocentric vision. The statistics of infants' visual experience guide referential processing as that high-frequency objects coincide with the object names normatively learned first (Clerkin, Hart, Rehg, Yu, & Smith, 2017). This and other resources to reduce referential ambiguity such as eye movements and ostensive display (Smith et al., 2010; Waxman & Gelman, 2009) are not accessible or highly limited to blind children. In these conditions, sound symbolic cues may play an important role in solving referential ambiguity (Imai & Kita, 2014) scaffolding form-to-meaning mapping and word learning.

Acknowledgements

This work was supported by the European Research Council starting grant (MADVIS grant #337573) attributed to OC. We are also extremely thankful to our participants and the Unione Ciechi e Ipovedenti in Milano and Trento, Savona, Trieste and Cuneo.

Author's contribution

RB, MB and OC design the study; RB and MB analyzed the data; RB, MB and OC wrote the paper.

Data availability

The stimuli of experiment 2 and all the data and analysis code are available at https://osf.io/yc4ud/.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2019.01.006.

---

[1] It is worth noticing that, in a less conservative model without random slopes and only random intercepts (Matuschek, Kliegl, Vasishth, Baayen, & Bates, 2017), the interaction Voicing*Group becomes significant.

References

Ahlner, F., & Zlatev, J. (2010). Cross-modal iconicity: A cognitive semiotic approach tosound symbolism. *Sign Systems Studies, 38*(1), 298–348.

Aveyard, M. E. (2012). Some consonants sound curvy: Effects of sound symbolism onobject recognition. *Memory & Cognition, 40*(1), 83–92. https://doi.org/10.3758/s13421-011-0139-3.

Bar, M., & Neta, M. (2006). Humans prefer curved visual objects. *Psychological Science,17*(8), 645–648. https://doi.org/10.1111/j.1467-9280.2006.01759.x.

Barilari, M., de Heering, A., Crollen, V., Collignon, O., & Bottini, R. (2018). Is red heavierthan yellow even for blind? *I-Perception, 9*(1), https://doi.org/10.1177/2041669518759123 204166951875912.

Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Psychology, 4*(December), 1–2. https://doi.org/10.3389/fpsyg.2013.00328.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (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

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects modelsusing lme4. *Eprint arXiv:1406.5823,* 51.

Blasi, D. E., Wichmann, S., Hammarström, H., Stadler, P. F., & Christiansen, M. H. (2016). Sound–meaning association biases evidenced across thousands of languages. *Pnas, 113*(39), 10818–10823. https://doi.org/10.1073/pnas.1605782113.

Bremner, A. J., Caparos, S., Davidoff, J., de Fockert, J., Linnell, K. J., & Spence, C. (2013). "Bouba" and "Kiki" in Namibia? A remote culture make similar shape-sound matches, but different shape-taste matches to Westerners. *Cognition, 126*(2), 165–172. https://doi.org/10.1016/j.cognition.2012.09.007.

Chandran, K. S., Banerjee, S., & Ghosh, K. (2017). A study on crossmodal correspondence in sensory pathways through forced choice task and frequency based correlation in sound-symbolism. *International Conference on Mining Intelligence and Knowledge Exploration (MIKE)* (pp. 212–220). . http://doi.org/10.1007/978-3-319-71928-3.

Chen, Y.-C., Huang, P.-C., Woods, A., & Spence, C. (2016). When "Bouba" equals "Kiki": Cultural commonalities and cultural differences in sound-shape correspondences. *Scientific Reports, 6*(1), 26681. https://doi.org/10.1038/srep26681.

Chéreau, C., Gaskell, M. G., & Dumay, N. (2007). Reading spoken words: Orthographic effects in auditory priming. *Cognition, 102*(3), 341–360. https://doi.org/10.1016/j.cognition.2006.01.001.

Clerkin, E. M., Hart, E., Rehg, J. M., Yu, C., & Smith, L. B. (2017). Real-world visual statistics and infants' first-learned object names. *Philosophical Transactions of the Royal Society B: Biological Sciences, 372*(1711), 20160055. https://doi.org/10.1098/rstb.2016.0055.

Cuskley, C., Simner, J., & Kirby, S. (2015). Phonological and orthographic influences in the bouba – kiki effect. *Psychological Research, 81*(1), 119–130. https://doi.org/10.1007/s00426-015-0709-2.

D'Onofrio, A. (2014). Phonetic detail and dimensionality in sound-shape correspondences: Refining the Bouba-Kiki Paradigm. *Language and Speech, 57*(3), 367–393. https://doi.org/10.1177/0023830913507694.

de Saussure, F. (1959). *Course of general linguistics.* New York: Philosophical Library.

Denys, K., Vanduffel, W., Fize, D., Nelissen, K., Peuskens, H., Van Essen, D. C., & Orban, G. A. (2004). The processing of visual shape in the cerebral cortex of human and nonhuman primates: A functional magnetic resonance imaging study. *Journal of Neuroscience, 24*(10), 2551–2565. https://doi.org/10.1523/JNEUROSCI.3569-03.2004.

Deroy, O., Fasiello, I., Hayward, V., & Auvray, M. (2016). Differentiated audio-tactile correspondences in sighted and blind individuals. *Journal of Experimental Psychology: Human Perception and Performance, 42*(8), 1204–1214. https://doi.org/10.1037/xhp0000152.

Deroy, O., Fasiello, I., Hayward, V., Auvray, M., Deroy, O., Fasiello, I., & Hayward, V. (2016). Differentiated audio-tactile correspondences in sighted and blind individuals. *Journal of Experimental Psychology: Human Perception and Performance, 42*(8), 1204–1214. https://doi.org/10.1037/xhp0000152.

Deroy, O., & Spence, C. (2013). Are we all born synaesthetic? Examining the neonatal synaesthesia hypothesis. *Neuroscience and Biobehavioral Reviews, 37*(7), 1240–1253. https://doi.org/10.1016/j.neubiorev.2013.04.001.

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. *Trends in Cognitive Sciences, 19*(10), 603–615. https://doi.org/10.1016/j.tics.2015.07.013.

Doyle, J. R., & Bottomley, P. A. (2011). Mixed messages in brand names: Separating the impacts of letter shape from sound symbolism. *Psychology & Marketing, 28*(7), 749–762. https://doi.org/10.1002/mar.

Fort, M., Martin, A., & Peperkamp, S. (2015). Consonants are more important than vowels in the bouba-kiki effect. *Language and Speech,* 122–140. https://doi.org/10.1177/0023830914534951.

Fryer, L., Freeman, J., & Pring, L. (2014). Touching words is not enough: How visual experience influences haptic-auditory associations in the "Bouba-Kiki" effect. *Cognition, 132*(2), 164–173. https://doi.org/10.1016/j.cognition.2014.03.015.

Hamilton-Fletcher, G., Pisanski, K., Reby, D., Stefańczyk, M., Ward, J., & Sorokowska, A. (2018). The role of visual experience in the emergence of cross-modal correspondences. *Cognition, 175*(February), 114–121. https://doi.org/10.1016/j.cognition.2018.02.023.

Hung, S.-M., Styles, S. J., & Hsieh, P.-J. (2017). Can a word sound like a shape before you have seen it? Sound-shape mapping prior to conscious awareness. 95679761667731 *Psychological Science, 28*(3), https://doi.org/10.1177/0956797616677313.

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

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

Imai, M., Miyazaki, M., Yeung, H. H., Hidaka, S., Kantartzis, K., Okada, H., & Kita, S. (2015). Sound symbolism facilitates word learning in 14-month-olds. *PLoS ONE, 10*(2), 1–17. https://doi.org/10.1371/journal.pone.0116494.

JASP Team. (2017). JASP (Version 0.8.1.2) [Computer software].

Jeffreys, H. (1961). *Theory of probability.* New York: Oxford University Press.

Kantartzis, K., Imai, M., & Kita, S. (2011). Japanese sound-symbolism facilitates word learning in English-speaking children. *Cognitive Science, 35*(3), 575–586. https://doi.org/10.1111/j.1551-6709.2010.01169.x.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association, 90*(430), 773–795.

Kirby, J. P., & Ladd, D. R. (2016). Effects of obstruent voicing on vowel *F* 0: Evidence from "true voicing" languages. *The Journal of the Acoustical Society of America, 140*(4),

Knoeferle, K., Li, J., Maggioni, E., & Spence, C. (2017). What drives sound symbolism? Different acoustic cues underlie sound-size and sound-shape mappings. *Scientific Reports, 7*(1), 5562. https://doi.org/10.1038/s41598-017-05965-y.

Kohler, W. (1947). *Gestalt psychology* (2nd ed.). New York, NY: Liveright Publishing.

Lockwood, G., Dingemanse, M., & Hagoort, P. (2016). Sound-symbolism boosts novel word learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(8), 1274–1281. https://doi.org/10.1037/a0028642.

Mathôt, S., Grainger, J., & Strijkers, K. (2017). Pupillary responses to words that convey a sense of brightness or darkness. *Psychological Science,* 1–9. https://doi.org/10.1177/0956797617702699.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language, 94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001.

Maurer, D., & Mondloch, C. J. (2005). Neonatal Synesthesia: a Reevaluation. Synesthesia: Perspectives from Cognitive Neuroscience. Retrieved from https://login.proxy.library.emory.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2005-09449-010&site=ehost-live.

Maurer, D., Pathman, T., & Mondloch, C. J. (2006). The shape of boubas: Sound-shape correspondences in toddlers and adults. *Developmental Science, 9*(3), 316–322. https://doi.org/10.1111/j.1467-7687.2006.00495.x.

McCormick, K., Kim, J. Y., List, S., & Nygaard, L. C. (2015). Sound to meaning mappings in the bouba – kiki effect. *CogSci, 2015*, 1565–1570.

Monaghan, P., Mattock, K., & Walker, P. (2012). The role of sound symbolism in language learning. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 38*(5), 1152–1164. https://doi.org/10.1037/a0027747.

Nielsen, A. K. S., & Rendall, D. (2013). Parsing the role of consonants versus vowels in the classic Takete-Maluma phenomenon. *Canadian Journal of Experimental Psychology, 67*(2), 153–163. https://doi.org/10.1037/a0030553.

Nielsen, A., & Rendall, D. (2011). The sound of round: Evaluating the sound-symbolic role of consonants in the classic Takete-Maluma phenomenon. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale, 65*(2), 115–124. https://doi.org/10.1037/a0022268.

Nygaard, L. C., Cook, A. E., & Namy, L. L. (2009). Sound to meaning correspondences facilitate word learning. *Cognition, 112*(1), 181–186. https://doi.org/10.1016/j.cognition.2009.04.001.

Ohala, J. J. (1997). Sound Symbolism. 4th Seoul International Conference on Linguistics [SICOL], (Ching 1982), 98–103. http://doi.org/10.1525/jlin.1996.6.1.109.

Ozturk, O., Krehm, M., & Vouloumanos, A. (2012). Sound symbolism in infancy: Evidence for sound-shape cross-modal correspondences in 4-month-olds. *Journal of Experimental Child Psychology, 114*(2), 173–186. https://doi.org/10.1016/j.jecp.2012.05.004.

Parise, C. V., & Spence, C. (2012). Audiovisual crossmodal correspondences and sound symbolism: A study using the implicit association test. *Experimental Brain Research, 220*(3–4), 319–333. https://doi.org/10.1007/s00221-012-3140-6.

Perlman, M., Dale, R., & Lupyan, G. (2015). Iconicity can ground the creation of vocal symbols. 150152 *Royal Society Open Science, 2*(8), https://doi.org/10.1098/rsos.150152.

R Core Team. (2013). R: A language and environment for statistical computing.

Ramachandran, V. S., & Hubbard, E. H. (2001). Synaesthesia—A window into perception, thought and language. *Journal of Consciousness Studies, 8*(12), 3–34. https://doi.org/10.1111/1468-0068.00363.

Shang, N., & Styles, S. J. (2017). Is a high tone pointy? Speakers of different languages match Mandarin Chinese tones to visual shapes differently. *Frontiers in Psychology, 8*(DEC), 1–13. https://doi.org/10.3389/fpsyg.2017.02139.

Sidhu, D. M., & Pexman, P. M. (2017). Five mechanisms of sound symbolic association. *Psychonomic Bulletin and Review, 1–25*. https://doi.org/10.3758/s13423-017-1361-1.

Sidhu, D. M., Pexman, P. M., & Saint-Aubin, J. (2016). From the Bob/Kirk effect to the Benoit/Éric effect: Testing the mechanism of name sound symbolism in two languages. *Acta Psychologica, 169*, 88–99. https://doi.org/10.1016/j.actpsy.2016.05.011.

Smith, L. B., Colunga, E., & Yoshida, H. (2010). Knowledge as process: Contextually cued attention and early word learning. *Cognitive Science, 34*(7), 1287–1314. https://doi.org/10.1111/j.1551-6709.2010.01130.x.

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics, 73*(4), 971–995. https://doi.org/10.3758/s13414-010-0073-7.

Styles, S. J., & Gawne, L. (2017). When does maluma/takete fail? Two key failures and a meta-analysis suggest that phonology and phonotactics matter. 204166951772480 *I-Perception, 8*(4), https://doi.org/10.1177/2041669517724807.

Turoman, N., & Styles, S. J. (2017). Glyph guessing for "oo" and "ee": Spatial frequency information in sound symbolic matching for ancient and unfamiliar scripts. *Royal*

Society Open Science, 4(9), https://doi.org/10.1098/rsos.170882.

Waxman, S. R., & Gelman, S. A. (2009). Early word-learning entails reference, not merely associations. *Trends in Cognitive Sciences, 13*(6), 258–263. https://doi.org/10.1016/j.tics.2009.03.006.

Westbury, C. (2005). Implicit sound symbolism in lexical access: Evidence from an interference task. *Brain and Language, 93*(1), 10–19. https://doi.org/10.1016/j.bandl.2004.07.006.

Westbury, C., Hollis, G., Sidhu, D. M., & Pexman, P. M. (2017). Weighing up the evidencefor sound symbolism: Distributional properties predict cue strength. *Journal of Memory and Language*. https://doi.org/10.1016/J.JML.2017.09.006.

Ziegler, J. C., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin & Review, 5*(4),683–689. https://doi.org/10.3758/BF03208845.