Sound symbolism is modulated by linguistic experience

João Delgado[1,2], Rodrigo Pereira[2], Miguel F. Ferreira[3,4], António Farinha-Fernandes[1], José C. Guerreiro[1], Bruno Faustino[1], Miguel Domingues[1], Paulo Ventura[1]

[1]CICPSI, Faculdade de Psicologia, Universidade de Lisboa
[2]CLUL, Faculdade de Letras, Universidade de Lisboa
[3]Instituto de Medicina Molecular João Lobo Antunes, Faculdade de Medicina, Universidade de Lisboa
[4]Instituto de Farmacologia e Neurociências, Faculdade de Medicina, Universidade de Lisboa

Abstract:

The bouba-kiki effect, or the symbolic association between certain speech sounds and rounded or angular shapes, is widely thought to be universal. However, two studies have failed to replicate this effect with neurotypical participants in the classical paradigm, one conducted in Papua New Guinea (Ross & Rogers, 1975), and the other conducted in Nepal (Styles & Gawne, 2017). As both experiments employed auditory stimuli inconsistent with the sound structure of the respective native language, Styles and Gawne (2017) proposed that pseudoword legality is a prerequisite for sound-symbolic associations to form. In this study, we conducted the first experimental test of this hypothesis, by assessing participants' performance on the bouba-kiki task as a function of pseudoword phonotactic legality. Our results indicate that phonotactic violations may disrupt the bouba-kiki effect, albeit only when they cause the speech stimuli to be perceived as significantly *strange* (not "word-like"). We thus conclude that sound symbolism fails whenever phonotactic violations prevent the assemblance of the phonological representations of the target pseudowords.

Keywords: Sound symbolism, bouba-kiki effect, phonotactic legality, phonological representations, European Portuguese

Palavras-chave: Simbolismo sonoro, efeito bouba-kiki, legalidade fonotática, representações fonológicas, Português Europeu

1. Introduction

Sound symbolism, or phonetic symbolism, refers to the non-arbitrary mapping between speech sounds and meaning. A good illustration of this is the bouba-kiki phenomenon[1] (Ramachandran & Hubbard, 2001), or the tendency to associate pseudowords like "bouba" to rounded shapes and pseudowords like "kiki" to angular shapes. Sound-symbolic associations with size, brightness, strength, speed, color and taste, among others, have also been reported (see Lockwood & Dingemanse, 2015 and references therein).

Until recently, sound symbolism was generally regarded as a curiosity of little significance, given that in natural languages the relationship between word forms and word meanings is largely arbitrary (Saussure, 1959). However, since the publication of Ramachandran and Hubbard's (2001) influential paper, there has been increasing recognition of the theoretical value of sound symbolism (Imai & Kita, 2014; Nielsen & Dingemanse, 2020), as well as of its pervasiveness in natural languages (Blasi et al., 2016). In fact, it has been argued that sound symbolism may have been instrumental in the emergence of language, and that it may be a fundamental force driving language development.

This view, known as the "sound symbolism bootstrapping hypothesis", was systematized by Imai and Kita (2014), who propose that sound symbolism derives from a biologically endowed ability to form cross-modal

[1] Also known as takete-maluma phenomenon (Kohler, 1929, 1947).

sensorial maps[2] (following Ramachandran & Hubbard, 2001) which, in turn, promote referential insight (i.e., the realization that speech sounds may function as signifiers) and facilitate the establishment of particular word-referent associations. In accordance with this account, recent studies showed that sound symbolism facilitates the acquisition of new words by children (see Imai & Kita, 2014 and references therein), whose early vocabularies reveal a high prevalence of sound symbolic words (Monaghan et al., 2014).

Proponents of the sound symbolism bootstrapping hypothesis often point to empirical research on the bouba-kiki effect as support for the assumption that (at least some) sound-symbolic associations may be universal. In the classical bouba-kiki paradigm, participants are simultaneously presented with two shapes, one rounded and the other angular/spiky, side by side, and two auditory pseudowords[3], one including "round sounding" phonemes, such as the sonorant consonants /m/ and /l/, the voiced plosive /b/ and the rounded back vowels /u/ and /o/, and the other including "angular/spiky sounding" phonemes, such as the voiceless plosives /t/ and /k/ and the unrounded front vowels /i/ and /e/. Their task consists in matching the pseudowords and the shapes, intuitively. The emergent effect is quite strong: according to a recent meta-analysis (Styles & Gawne, 2017), when the auditory stimuli are composed primarily by the phonemes listed above, 84% to 94% of people agree that "round sounding" pseudowords, in the sense specified above, go together with rounded shapes, whereas "angular/spiky sounding" pseudowords go together with angular/spiky shapes. Importantly for the present discussion, this effect has been found with populations from different linguistic and cultural contexts (see Lockwood & Dingemanse, 2015 for a review), including the Himba (Bremner et al., 2013), a remote, non-western, illiterate Namibian population, thus providing support for the assumption of universality. Studies with children point in the same direction, revealing that both toddlers and infants are sensitive to the bouba-kiki phenomenon (Asano et al., 2015; Fort et al., 2018; Maurer et al., 2006; Ozturk, Krehm & Vouloumanos, 2013; cf. Fort et al., 2013; Pejovic & Molnar, 2017).

Despite the well attested robustness of the bouba-kiki effect, there are two studies in the literature reporting failed replication attempts with neurotypical participants in the classical paradigm[4], one testing speakers of Hunjara, from Papua New Guinea (Ross & Rogers, 1975), and the other testing speakers of Syuba, from Nepal (Styles & Gawne, 2017). In both cases, performance on the bouba-kiki task was at chance level, indicating that participants did not form sound-symbolic associations. Under the assumption of universality, these null results remain to be clarified.

Styles and Gawne (2017) present the first attempt of specifying the conditions under which the bouba-kiki effect might fail. The authors propose that sound symbolism may be modulated by linguistic experience, such that the typical mapping between pseudowords and meanings may be disrupted when the pseudowords do not conform with the phonology and phonotactics of the participants' language, i.e., when the pseudowords are not "word-like" – henceforth referred to as "invalid pseudowords". This view is supported by the observation that both studies that yielded a null result in the classical bouba-kiki paradigm employed auditory stimuli that did not match the sound structure of participants' languages[5]: Ross & Rogers (1975) used the stimuli "takete" and

[2] Some cross-modal mappings may only be accessible to humans: Margiotoudi and colleagues (2019) failed to elicit the bouba-kiki effect in a sample of non-human primates, suggesting that it may depend on brain structures/connections specific to Homo Sapiens (e.g., a strongly developed Arcuate Fasciculus - see Margiotoudi et al., 2019 for further discussion).

[3] In a common variant of the task, only one pseudoword per trial is presented (e.g., Fort, Martin & Peperkamp, 2015).

[4] There is evidence that the bouba-kiki effect is disrupted in autism (Oberman & Ramachandran, 2008; Occeli et al., 2013) and dyslexia (Drijvers, Zaadnoordijk & Dingemanse, 2015).

[5] It should be noted that there is one study whose results seem to go against Styles and Gawne's (2017) prediction, namely that by Bremner and colleagues (2013), who found the bouba-kiki effect even though one of the pseudowords used, "bouba", included a sound, [b], which is unattested in the native language of the participants (Otjiherero). One should be careful, however, in attaching too much theoretical significance to this result; as noted by Styles and Gawne (2017), even though [b] is not part of the phonological repertoire of Otjiherero, the closely related phone [p] is, raising the possibility that participants perceptually assimilated [b] to [p] (Best, McRoberts & Goodell, 2001). Given that participants were asked to articulate the pseudowords before completing the task, this hypothesis is rendered especially plausible.

"maluma", which included two sounds unattested in the language spoken by the participants (Hunjara), [t^h] and [l], respectively; and Styles and Gawne (2017) used the stimuli "bubu" and "kiki", which included combinations of sounds unattested in the language spoken by the participants (Syuba), namely a word-medial [k^h], in "kiki", and a word-final [u], in "bubu", and which involved reduplicated tones, rather than the tone levelling that normally occurs in long Syuba words (given the unusual phonotactic structure of the stimuli, the native who recorded the two tokens used in the experiment pronounced the tones in each pseudoword as though it was made of two first syllables).

The hypothesis formulated by Styles and Gawne (2017) has the merit of generating testable predictions. If sound symbolism is disrupted by phonological and phonotactic/tonotactic violations, one would expect to find the bouba-kiki effect in speakers of a given language if the stimuli used in the bouba-kiki task conform with that language's sound structure, but not otherwise.

In this study, we present the first experimental test of this hypothesis, by assessing the performance of European Portuguese (EP) speakers in the bouba-kiki task as a function of pseudoword phonotactic legality.

Testing Styles and Gawne's (2017) prediction seems relevant not just for the discussion of the universality of sound symbolism, but also for the debate on which level of analysis, i.e., phonological, articulatory or acoustic[6], is best suited to account for sound symbolism. Empirical studies employing carefully controlled auditory stimuli have found angular and rounded shapes to be differentially related to particular phonetic features of consonants and vowels (D'Onofrio, 2014; Fort, Martin & Peperkamp, 2015), particular acoustic properties of vowels (Knoeferle et al., 2017), and particular phonemes (Monaghan & Fletcher, 2019; Westbury et al., 2018). However, since differences in phonetic features (which combine to form phonemes) imply acoustic and articulatory differences, it is hard to disentangle, on the basis of these data, whether the effect is driven by phonological, acoustic or articulatory representations. The test of whether phonotactics interfere with the bouba-kiki phenomenon, on the other hand, allows us to circumvent this problem: if the phonological context in which particular phonemes occur, i.e., valid or invalid sequence of phonemes, determines whether the bouba-kiki effect is preserved or disrupted, one may safely conclude that sound symbolism is mediated by phonological representations.

We report three experiments. In Experiments 1 and 3, we explored whether phonotactic violations disrupt the bouba-kiki effect in a classic bouba-kiki task. In Experiment 2, we ran a classification task aimed at quantifying the strength of the phonotactic violations employed in Experiments 1 and 3. Participants were independently recruited for each experiment.

2. Experiment 1

The goal of Experiment 1 was to assess whether pseudoword legality is a prerequisite for the bouba-kiki effect. We therefore compared the performance of two groups of participants on a two-trial bouba-kiki task, one hearing phonotactically legal pseudowords, and another hearing phonotactically illegal pseudowords. If phonotactic violations disrupt the bouba-kiki effect, we should observe, on both trials, a significantly lower, potentially at chance, effect on the latter group.

2.1. Participants

[6] There is some evidence that orthography may also play a role (Cuskley, Simner & Kirby, 2017). However, given that the bouba-kiki effect has been found with pre-literate children (Fort et al., 2018) and illiterate adults (Bremner et al., 2013), it is reasonable to assume that orthographic influences are minor.

Seventy-two healthy adults (65 females) participated in the experiment. All were native speakers of European Portuguese, with no language disability, and had normal or corrected to normal vision. Participants' ages ranged from 17 to 41 (M = 19.6; SD = 4.3).

2.2. Materials

For the auditory stimuli, we created four valid pseudowords and four invalid pseudowords, the former being possible words in EP, and the latter being words that violate the phonotactic rules of EP. The pseudowords were created in pairs. First, we created the two pairs of valid pseudowords /pi.ki.ˈtɛ/-/ma.ˈno.bu/ and /ki.ˈpɛ.ki/-/ba.ˈlu.ba/, with each pair including a round sounding pseudoword, i.e., containing canonical round sounding phonemes (/m, l, b, u, o/), and an angular sounding pseudoword, i.e., containing canonical angular sounding phonemes (/t, k, i, ɛ/). Then, from these stimuli, we created the two pairs of invalid pseudowords /pi.tki.ˈɛ/-/man.ˈbo.u/ and /ˈki.pɛ.ki/-/ˈbaul.ba/, by changing the order of phonemes or the stress pattern of the valid pseudowords. Pairs of valid and invalid pseudowords were thus matched for phonemic constitution of the pseudowords. See Table 1 for a summary of the stimuli used in Exp.1, including the phonetic transcriptions of the pseudowords, a description of the valid-to-invalid conversion processes and a description of the phonotactic violations in the invalid pseudowords. The auditory stimuli were recorded in an anechoic chamber by a trained linguist, and their peak volume was standardized in AudacityTeam (http://audacityteam.org/).

Symbolic Group	Valid pseudoword	Invalid Pseudoword	Valid – Invalid conversion process	Phonotactic violation (Mateus et al. 2016; Vigário & Falé, 1994)
Angular	/ki.ˈpɛ.ki/ [ki.ˈpɛ.ki]	/ˈki.pɛ.ki/ [ˈki.pɛ.ki]	Transposition of the main stress to the first syllable	No final /i/ vowel in EP in a proparoxytone word
	/pi.ki.ˈtɛ/ [pi.ki.ˈtɛ]	/pi.tki.ˈɛ/ [pi.tki.ˈɛ]	Metathesis of [t] to the second syllable	Inexistence of the phonological /tk/ consonant sequence
Rounded	/ba.ˈlu.ba/ [bɐ.ˈlu.bɐ]	/ˈbaul.ba/ [ˈbawɫ.bɐ]	Metathesis of /l/ to the syllable coda of the first syllable	Inexistence of /l/ in syllable coda after a diphthong [aw][7]
	/ma.ˈno.bu/ [mɐ.ˈno.bu]	/man.ˈbo.u/ [mɐn.ˈbo.u]	Metathesis of [n] to the coda of the first syllable	Inexistence of the consonant sequence /nb/ with /n/ in coda

Table 1. Summary of the auditory stimuli used in Experiment 1

The visual stimuli consisted of two pairs of shapes, each including a rounded and an angular shape, side by side. One of the pairs was the stimuli used in Styles and Gawne (2017)[8] (see figure 1), and the other pair was constructed in Microsoft Visio (see Figure 2). All the shapes were painted in black and occupied the same area size on the screen.

[7] This violation is related to the maximum syllable weight allowed in Portuguese.
[8] Downloaded from the Open Science Framework repository - https://osf.io/wt95v/.

Figure 1. Pair of bouba-kiki shapes (taken from Styles & Gawne, 2017)

Figure 2. Pair of bouba-kiki shapes (constructed in Microsoft Visio)

2.3. Procedure

We constructed two versions of the bouba-kiki task, in Microsoft PowerPoint, with version 1 including the valid pseudowords, and version 2 including the invalid pseudowords. Participants were randomly allocated to one of the two versions, with half of the participants (n = 36) performing version 1, and the other half performing version 2. The bouba-kiki task consisted of two trials (2 PowerPoint slides). In the first trial, participants heard the valid pseudowords /pi.ki.ˈtɛ/-/ma.ˈno.bu/ or the invalid pseudowords /pi.tki.ˈɛ/-/man.ˈbo.u/, depending on the task version, and saw the pair of shapes illustrated in Figure 1. In the second trial, participants heard the valid pseudowords /ki.ˈpɛ.ki/-/ba.ˈlu.ba/ or the invalid pseudowords /ˈki.pɛ.ki/-/ˈbaul.ba/, depending on the task version, and saw the pair of shapes illustrated in Figure 2. The pseudowords were always played through the same headphones and the shapes were always presented on the same screen. Sub-versions of the task were constructed so as to counterbalance the order of presentation of the pseudowords and shapes in each trial.

At the start of Experiment 1, participants were informed that they would hear two strange sounds and see two strange figures in each trial. They were told that the sounds were Martian names for the figures, such that each sound corresponded to one figure, and that their task consisted in pointing to the figure that they thought the first sound corresponded to. It was stressed that there were no right or wrong answers and that we were only interested in participants' intuitions. Participants answered under no time constraints, and the auditory stimuli could be replayed as many times as participants wished. The answers were registered by an experimenter.

2.4. Results

Figure 3 shows the percentage of expected associations for each pair of pseudowords, i.e., associations of pseudowords including angular and round sounding phonemes with angular and rounded shapes, respectively. We used one-way chi-square tests to assess whether the number of expected associations for a given pair of auditory stimuli differed from what would be expected by chance, and two-way chi-square tests to assess whether the number of expected associations for each pair of valid pseudowords was significantly different from the number of expected associations for its phonemically matched pair of invalid pseudowords.

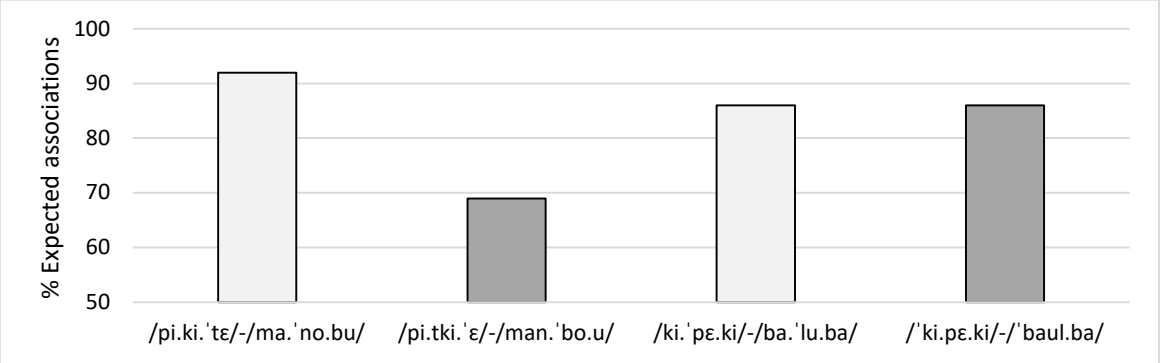

Figure 3. Percentage of expected associations for each pair of auditory stimuli. Light bars correspond to valid pseudowords, whereas dark bars correspond to invalid pseudowords

Performance was above-chance for all pseudoword pairs (/pi.kiˈtɛ/-/maˈno.bu/, $X^2(1, N=36) = 25$, $p < .001$; /pi.tkiˈɛ/-/manˈbo.u/, $X^2(1, N=36) = 5.44$, $p = .020$; /kiˈpɛ.ki/-/baˈlu.ba/, $X^2(1, N=36) = 18.78$, $p < .001$; /ˈki.pɛ.ki/-/ˈbaul.ba/, $X^2(1, N=36) = 18.78$, $p < .001$). The number of expected associations was significantly higher for the valid pseudowords /pi.kiˈtɛ/-/maˈno.bu/ than for the invalid pseudowords /pi.tkiˈɛ/-/manˈbo.u/, $X^2(1, N=72) = 5.68$, $p = .017$. However, the number of expected associations for the valid pseudowords /kiˈpɛ.ki/-/baˈlu.ba/ was not significantly different from the number of expected associations for the invalid pseudowords /ˈki.pɛ.ki/-/ˈbaul.ba/, $X^2(1, N=72) = 0$, $p = 1$.

2.5. Discussion

Both pairs of valid pseudowords elicited a bouba-kiki effect, as expected. However, the evidence for an effect of phonotactics on the bouba-kiki effect was mixed. Phonotactic illegality disrupted the bouba-kiki effect in only one of the two pairs of illegal stimuli (/pi.tkiˈɛ/-/manˈbo.u/). Furthermore, the disruption was only partial, as the bouba-kiki effect still emerged, albeit in a reduced form. Assuming Styles and Gawne's (2017) hypothesis, the partial disruption might be explained by EP's optional elision of [ɨ], which allows for long strings of consonants to form, like in /desprɵteʒɐr/, realized phonetically as [dʃprutˈʒɐr], with the unstressed /e/ being ellided: it is possible that some participants inserted a /e/ in the phonological representation of /pi.tkiˈɛ/ and /manˈbo.u/, after /t/ and /n/, respectively, circumventing the phonological violations, and therefore forming the expected sound-symbolic associations. The different behavior of the other pair of illegal pseudowords, on the other hand, is harder to explain. However, a slightly modified version of Styles and Gawne's (2017) proposal may accommodate our data: retaining the idea that "wordiness" is the critical factor determining sound symbolism success or failure, we propose that only some phonotactic violations are sufficiently strong to disrupt typical phonological processing. In other words, not all descriptively invalid pseudowords are perceived as significantly strange (not "word-like"). From this perspective, one may predict that the phonotactic violations

of the invalid pseudowords that disrupted the bouba-kiki effect are stronger than the phonotactic violations of the invalid pseudowords that did not disrupt the effect. In order to test this prediction, we ran a second experiment, aimed at quantifying the strength of the phonotactic violations of the invalid pseudowords used in Experiment 1.

3. Experiment 2

The goal of Experiment 2 was to assess whether the phonotactic violations that disrupted the bouba-kiki effect in Experiment 1 are stronger than those that didn't. We operationalized "strength of a phonotactic violation" as the perceived strangeness (distance from EP) of a pseudoword containing that violation. If the different behavior of the two pairs of invalid pseudowords used in Experiment 1 is due to different strengths of their phonotactic violations, the difference between valid and invalid pseudowords in perceived strangeness should map onto the difference between valid and invalid pseudowords in number of expected associations on the bouba-kiki task (Experiment 1).

3.1. Participants

Fifteen healthy adults (10 females) participated in the experiment. All were native speakers of European Portuguese, with no language disability, and had normal or corrected to normal vision. Participants' ages ranged from 18 to 23 (M = 20.1; SD = 1.6).

3.2. Materials

The auditory stimuli for Experiment 2 consisted of the eight pseudowords used in Experiment 1 and twelve additional pseudoword fillers. Half the fillers were valid pseudowords and half were invalid pseudowords, matched in terms of phonemic constitution. The pseudoword fillers were recorded by a trained linguist in an anechoic chamber, and the resulting audio files were normalized for peak amplitude in AudacityTeam (http://audacityteam.org/). Two presentation lists were created, each including the twenty pseudowords in a random order.

3.3. Procedure

The presentation lists were embedded in PowerPoint presentations. Eight participants saw one list, and the remaining saw the other. Each slide contained the label "sound 1…20" and a button that played a pseudoword. In the beginning of the experiment, participants were informed that they would hear nonsense names/nouns, some stranger than others, and were asked to classify the names/nouns in terms of the degree of perceived strangeness/distance from EP, in a scale of 1 to 9, wherein 1 would be a normal name/noun in EP and 9 something very distant from EP. Participants registered their answers in a spreadsheet.

3.4. Results

Figure 4 shows the mean classification of the pseudoword pairs used in Experiment 1 in perceived strangeness/distance from EP. We used Wilcoxon tests to assess whether the classification of each pair of valid pseudowords differed from the classification of its phonemically matched pair of invalid pseudowords. We found a significant difference between the valid pseudowords /pi.ki.ˈtɛ/-/ma.ˈno.bu/ and the invalid pseudowords /pi.tki.ˈɛ/-/man.ˈbo.u/, Z = -3.37, p = .001, but not between the valid pseudowords /ki.ˈpɛ.ki/-/ba.ˈlu.ba/ and the invalid pseudowords /ˈki.pɛ.ki/-/ˈbaul.ba/, Z = -.75, p = .452.

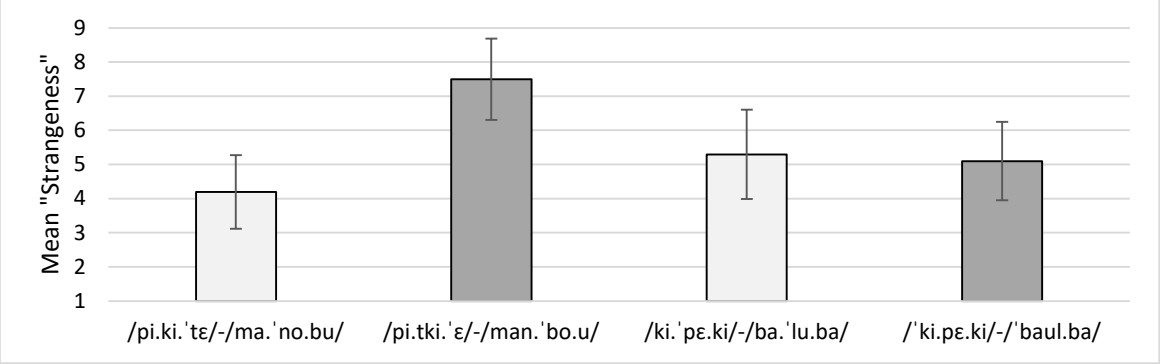

Figure 4. Mean classification in strangeness/distance from EP of the pseudoword pairs used in Experiment 1. Light bars correspond to valid pseudowords, whereas dark bars correspond to invalid pseudowords. Error bars represent standard deviations

3.5. Discussion

The results from the classification task were very clear: only the invalid pseudowords that disrupted the bouba-kiki effect in Experiment 1 were perceived as significantly deviant relative to their valid counterparts. This suggests that not all phonotactic violations, as defined by the inexistence of a particular sequence of sounds or stress pattern in a given language, disrupt phonological processing – both the stress violation in /ˈki.pɛ.ki/ (and, perhaps, stress violations more generally[9]) and the sound sequence violation in /ˈbaul.ba/ (see Table 1) may be among those violations which are not strong enough to disrupt phonological parsing[10]. We return to this point in the General Discussion.

Our results thus support the hypothesis that only some phonotactic violations (i.e., strong phonotactic violations) interfere with sound symbolism. Even so, strong conclusions are precluded by potential confounds in Experiment 1. For one, the stimuli that disrupted the bouba-kiki effect always occurred in the first trial, raising the possibility that the lack of disruption for the invalid pseudowords /ˈki.pɛ.ki/-/ˈbaul.ba/ was due to a change in strategy or response criteria from the first to the second trial. Secondly, as the disruption was only found for one pair of pseudowords, it is possible that it is not due to the presence of a strong phonotactic violation, per se, but to some idiosyncratic property of that particular pair of stimuli. We thus ran a third experiment, wherein we attempted to replicate, in a more controlled experimental setting, the effect found in Experiment 1, and generalize the effect of phonotactic illegality to a new pair of stimuli.

4. Experiment 3

The goal of Experiment 3 was to assess whether the effect of sound symbolism disruption found in Experiment 1 is reliable and whether it generalizes to a new pair of stimuli. Thus, half the pseudowords used in

[9] Possibly, a change in the main word stress is not a strong phonotactic violation in European Portuguese because the distinction is only observed in a few minimal-pairs.

[10] It is possible that, given the acoustic similarity between [ɬ] and the previous segment [w], participants failed to perceive the phonotactic violation in /ˈbaul.ba/, taking it as [ˈbaw.bɐ] or [ˈbaɬ.bɐ], which are both valid pseudowords. However, if the stress violation in /ˈki.pɛ.ki/ was sufficiently strong to disrupt phonological processing, an interference with the bouba-kiki effect would, presumably, still be expected, contrary to what we observed in Experiment 1.

the current experiment were taken from Experiment 1 - /pi.ki.ˈtɛ/-/ma.ˈno.bu/ and /pi.tki.ˈɛ/-/man.ˈbo.u/, - and half were new. We compared the performance of two groups of participants on a two-trial bouba-kiki task, one hearing phonotactically legal pseudowords, and another hearing pseudowords with strong phonotactic violations. If strong phonotactic violations disrupt the bouba-kiki effect, we should observe, on both trials, a significantly lower, potentially at chance, effect on the latter group.

4.1. Participants

One hundred and twelve healthy adults (72 females) participated in the experiment. All were native speakers of European Portuguese, with no language disability, and had normal or corrected to normal vision. Participants' ages ranged from 17 to 40 (M = 20.9; SD = 4.3).

4.2. Materials

The auditory stimuli were two pairs of valid pseudowords and two pairs of invalid pseudowords: the pairs /pi.ki.ˈtɛ/-/ma.ˈno.bu/ and /pi.tki.ˈɛ/-/man.ˈbo.u/, from Experiment 1, and the new pairs /ka.ˈke.te/-/mo.ˈɾɔ.ʃa/ (valid) and /e.a.ˈtkek/-/ɔ.ʃm.ˈɾo.a/ (invalid), used in the classification task (Experiment 2) as fillers. The latter pairs of pseudowords, like the former, included a pseudoword containing canonical round sounding phonemes and a pseudoword containing canonical angular sounding phonemes. Importantly, the new invalid pseudowords were judged to be significantly stranger than their valid counterparts (see Figure 5), Z = -3.42, p = .001.

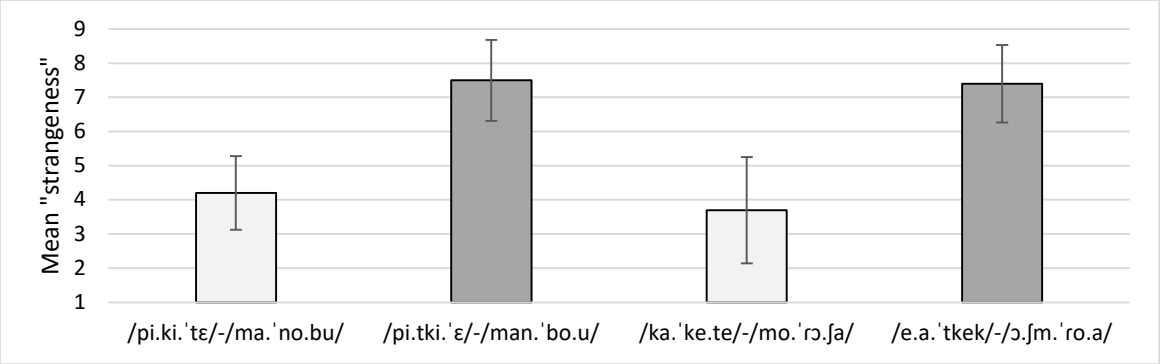

Figure 5. Mean classification in strangeness/distance from EP of the pseudoword pairs used in Experiment 3. Light bars correspond to valid pseudowords, whereas dark bars correspond to invalid pseudowords. Error bars represent standard deviations

See Table 2 for a summary of the new stimuli, including the phonetic transcriptions of the pseudowords, a description of the valid-to-invalid conversion processes and a description of the phonotactic violations in the invalid pseudowords.

Symbolic Group	Valid pseudoword	Invalid Pseudoword (non-word)	Valid – Invalid conversion process	Phonotactic violation (Mateus et al. 2016; Vigário & Falé, 1994)
Angular	/pi.ki.ˈtɛ/ [pi.ki.ˈtɛ]	/pi.tki.ˈɛ/ [pi.tki.ˈɛ]	Metathesis of [t] to the second syllable	Inexistence of the phonological /tk/ consonant sequence

	/ka.ˈke.te/ [kɐ.ˈke.tɨ]	/e.a.ˈtkek/ [ɨ.ɐ.ˈtkek]	Metathesis of various sounds	No [k] in coda, no [tk] consonant sequence and no initial [ɨ]
Rounded	/ma.ˈno.bu/ [mɐ.ˈno.bu]	/man.ˈbo.u/ [mɐn.ˈbo.u]	Metathesis of [n] to coda of the first syllable	Inexistence of the consonant sequence /nb/ with /n/ in coda
	/mo.ˈɾɔ.ʃa/ [mo.ˈɾɔ.ʃɐ]	/ɔ.ʃm.ˈɾo.a/ [ɔ.ʃm.ˈɾo.ɐ]	Metathesis of various sounds	Inexistence of [m] as a syllabic consonant, and /ɾ/ after a nasal

Table 2. Summary of the auditory stimuli used in Experiment 3

4.3. Procedure

We constructed two versions of the bouba-kiki task, in E-prime 2.0 software (Psychology Software Tools, Pittsburgh, PA), with version 1 including the valid pseudowords, and version 2 including the invalid pseudowords. Participants were randomly allocated to one of the two versions, with half of the participants (n = 56) performing version 1, and the other half performing version 2. The bouba-kiki task consisted of two trials. In one trial, participants heard the valid pseudowords /pi.ki.ˈtɛ/-/ma.ˈno.bu/ or the invalid pseudowords /pi.tki.ˈɛ/-/man.ˈbo.u/, depending on the task version, and saw the pair of shapes illustrated in Figure 2. In another trial, participants heard the valid pseudowords /ka.ˈke.te/-/mo.ˈɾɔ.ʃa/ or the invalid pseudowords /e.a.ˈtkek/-/ɔ.ʃm.ˈɾo.a/, depending on the task version, and saw the pair of shapes illustrated in Figure 1. The order of the trials was counterbalanced. The pseudowords were always played through the same headphones and the shapes were always presented on the same screen. The order of presentation of the pseudowords and shapes was counterbalanced.

The instructions were the same as in Experiment 1, except that participants were now asked to provide their answer by pressing a key.

4.4. Results

Figure 6 shows the percentage of expected associations for each pair of pseudowords, i.e., associations of pseudowords including angular and round sounding phonemes with angular and rounded shapes, respectively. We performed the same chi-square tests as in Experiment 1. Performance was at chance level for the invalid pseudowords /e.a.ˈtkek/-/ɔ.ʃm.ˈɾo.a/, X^2 (1, N=56) = 2.57, p = .109, and above-chance for all other pairs of pseudowords (/ka.ˈke.te/-/mo.ˈɾɔ.ʃa/, X^2 (1, N=56) = 10.29, p = .001, /pi.ki.ˈtɛ/-/ma.ˈno.bu/, X^2 (1, N=56) = 28.57, $p < .001$, /pi.tki.ˈɛ/-/man.ˈbo.u/, X^2 (1, N=56) = 5.79, $p < .016$). As in Experiment 1, we found a significant difference between the valid pseudowords /pi.ki.ˈtɛ/-/ma.ˈno.bu/ and the invalid pseudowords /pi.tki.ˈɛ/-/man.ˈbo.u/, X^2 (1, N=112) = 5.9, p = .015. However, the difference between the valid pseudowords /ka.ˈke.te/-/mo.ˈɾɔ.ʃa/ and the invalid pseudowords /e.a.ˈtkek/-/ɔ.ʃm.ˈɾo.a/ did not reach significance, X^2 (1, N=112) = 1.43, p = .231.

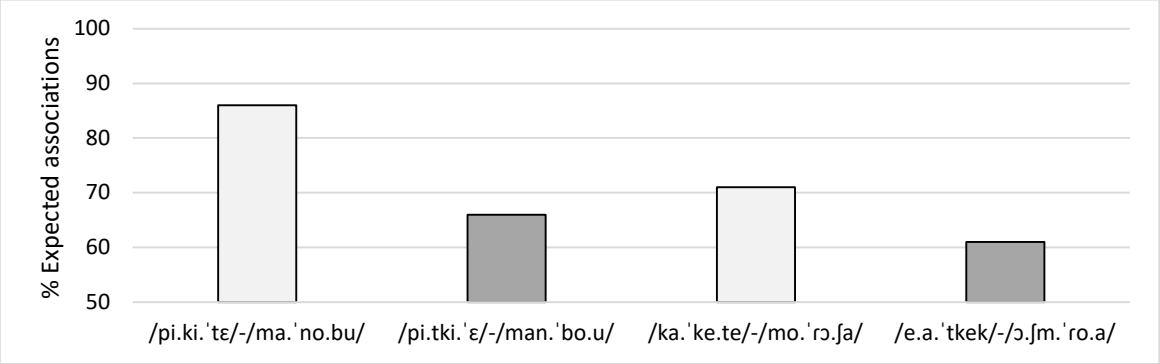

Figure 6. Percentage of expected associations for the pairs of pseudowords used in Experiment 3. Light bars correspond to valid pseudowords, whereas dark bars correspond to invalid pseudowords

4.5. Discussion

We replicated the relevant effect found in Experiment 1, namely a disruption of the bouba-kiki effect by the invalid pseudowords /pi.tki.ˈɛ/-/man.ˈbo.u/, thus ruling out the order of trials as a possible explanation for the different behavior of the two pairs of invalid pseudowords used in Experiment 1. These results provide further strength to the hypothesis that strong phonotactic violations interfere with sound symbolism. The data pertaining to the new stimuli are consistent with this interpretation, although less robust: even though the direct comparison between valid and invalid pseudowords did not emerge as statistically significant, the one-way chi-square tests suggest that only the performance with the invalid stimuli was at chance level. We propose that the failure to find a significant difference between the new valid and invalid pseudowords was due to an atypically low performance with the valid stimuli (71% expected associations, as opposed to ≥86% expected associations for the stimuli used in Experiment 1), which may be explained by the inclusion of less canonical round sounding phonemes (Styles & Gawne, 2017) - the less studied phonemes /ɾ, ʃ, ɔ,/, contained in /mo.ˈrɔ.ʃa/, may not be straightforwardly associated to rounded shapes, at least in EP. This somewhat atypical finding calls for more cross-linguistic research on the role of less studied phonemes on the bouba-kiki effect.

5. General Discussion

The experiments reported in this article show that the bouba-kiki effect is significantly reduced when the auditory stimuli are not perceived to be "word-like", in line with Styles and Gawne's (2017) suggestion. However, this effect is not determined by phonotactic illegality, per se; according to our data, whereas some phonotactic violations are sufficiently strong to disrupt phonological processing and, as a result, sound symbolism, others are not.

These results are important both for the debate on the universality of sound symbolism and for the debate on which level of analysis best accounts for sound symbolism. However, before we address these questions, a point on phonotactic legality and "wordiness" is in order. As noted, our data suggest that some phonotactically illegal pseudowords were not considered strange. In other words, it seems that some sequences of phonemes/stress patterns that do not occur in a language are nevertheless represented normally, as evidenced by the data from the classification task (Experiment 2). This is an interesting result, which suggests that experimental studies may feed important information to theories in Phonology. Why do some phonotactic violations render a pseudoword strange sounding, whereas others do not? Our data are too coarse to delve deeply

into these matters; nevertheless, we suggest that one plausible possibility is that the violation of *natural phonotactic constraints*, i.e., constraints that observe typologically common restrictions and/or are phonetically grounded, more strongly disrupts phonological processing than the violation of *unnatural phonotactic constraints*, i.e. constraints that do not meet the criteria for naturalness (Hayes & White, 2013, Prickett, 2018). In fact, Hayes and White (2013) found that participants rated words that violated natural phonotactic constraints of English as worse than words that violated unnatural phonotactic constraints (both natural and unnatural constraints were arrived at by the phonotactic learning system of Hayes & Wilson, 2008). This hypothesis may explain why the pairs /pi.tki.ˈɛ/-/man.ˈbo.u/ and /e.a.ˈtkek/-/ɔ.ʃm.ˈɾo.a/ were perceived as strange, whereas the pair /ˈki.pɛ.ki/-/ˈbaul.ba/ was not, as only the former included pseudowords violating a natural phonotactic constraint – failure to comply with the Sonority Principle (Selkirk, 1984; Clements, 1990).[11]

Much recent research supports the notion that the bouba-kiki effect, among other sound-symbolic associations, is universal (see Introduction). However, the failed replication attempts of Ross and Rogers (1975) and Styles and Gawne (2017), in conjunction with our own data, indicate that the bouba-kiki effect is mediated by high-level, phonological representations, which ought to conform with a speaker's knowledge of phonological and phonotactic constraints lest the effect may be disrupted. This suggests language-specificity in sound-symbolic associations. In our view, these two positions are not irreconcilable: sound symbolism may depend on a biologically-endowed ability to form cross-modal maps (perhaps by abstracting properties shared by the sound/articulation of speech and associated stimuli[12], Ramanchandran & Hubbard, 2001; see also Sidhu & Pexman, 2018) which, with accumulating linguistic experience, end up being mediated by phonological representations. If so, sound symbolism may function as a bootstrapping mechanism for language acquisition from the earlier stages, even though experience with a particular language ends up disallowing otherwise universal mappings between acoustic/articulatory representations and visual representations when the speech stimuli fail to be parsed into ordered sequences of phonemes (Styles & Gawne, 2017).

Interestingly, a recent study (Silva & Bellini-Leite, 2019) found the bouba-kiki effect with sine-wave speech both in a group of participants trained to perceive the stimuli as speech and in a group of participants who failed to discern speech sounds in the stimuli, even though the effect was stronger in the former (in a cross-modal matching task analogous to the classic bouba-kiki task). These results are somewhat at odds with our conclusions, as they suggest that both speech-specific and non-speech-specific representations may drive sound-symbolic associations. This could potentially explain why some of our invalid stimuli only partially disrupted the bouba-kiki effect: phonotactic illegality would compromise the speech-specific mechanism, but not the general perceptual mechanism. However, it would remain to be explained why one of our pairs of invalid pseudowords, /e.a.ˈtkek/-/ɔ.ʃm.ˈɾo.a/, as well as the stimuli used in Ross and Rogers (1975) and Styles and Gawne (2017), entailed a complete breakdown of sound symbolism. Future research could further explore whether sound symbolism is primarily driven by one or more mechanisms, and the conditions under which such mechanisms might be deployed.

Before closing the discussion, we would like to point out that, even though our data form a very compelling picture in light of Styles and Gawne's (2017) proposal, the low number of pseudowords used in our study demand caution in generalizing our conclusions. It would be important for future research to attempt to generalize our findings to larger, more diverse, sets of stimuli. Such endeavors could assess, in addition, whether

[11] An alternative possibility worthy of note is that the constraints violated by /ˈki.pɛ.ki/-/ˈbaul.ba/ are *accidentally true,* in the sense of Hayes and White (2013), i.e., they hold true of a language's lexicon but are not part of the phonotactic knowledge of native speakers. As applied to /ˈbaul.ba/, however, this seems unlikely, since the phonotactic violation follows from general principles concerning syllable weight in EP (but see Footnote 10).

[12] For instance, a speaker might realize that the abrupt burst of sound in the production of a voiceless stop consonant mimics the abrupt changes in direction in the outline of an angular shape, or that the lip rounding in the production of a rounded vowel mimics the rounded contour of a rounded shape, therefore predisposing him/her to associate words containing voiceless stops to angular shapes and words containing rounded vowels to rounded shapes (Ramachandran & Hubbard, 2001).

different kinds of phonotactic violations interfere with sound symbolism in different ways, and experimentally test Styles and Gawne's (2017) prediction that phonological violations (i.e., inclusion of sounds foreign to one's language) also disrupt the bouba-kiki effect. Furthermore, future research could investigate whether other language-specific factors, such as the frequency of co-occurrence of phonemes in the lexicon, also impact the bouba-kiki effect.

6. References

Asano, M., Imai, M., Kita, S., Kitajo, K., Okada, H., & Thierry, G. (2015) Sound symbolism scaffolds language development in preverbal infants. *Cortex*, *63*, 196-205.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001) Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *The Journal of the Acoustical Society of America*, *109*(2), 775-794.

Blasi, D. E., Wichmann, S., Hammarström, H., Stadler, P. F., & Christiansen, M. H. (2016) Sound–meaning association biases evidenced across thousands of languages. *Proceedings of the National Academy of Sciences*, *113*(39), 10818-10823.

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.

Clements, G. N. (1990) The role of the sonority cycle in core syllabification. In J. Kingston & M. Beckman (eds.) *Papers in Laboratory Phonology I*, 283-333. Cambridge: Cambridge University Press.

Cuskley, C., Simner, J., & Kirby, S. (2017) Phonological and orthographic influences in the bouba–kiki effect. *Psychological research*, *81*(1), 119-130.

D'Onofrio, A. (2014) Phonetic detail and dimensionality in sound-shape correspondences: Refining the bouba-kiki paradigm. *Language and Speech*, *57*(3), 367-393.

Drijvers, L., Zaadnoordijk, L., & Dingemanse, M. (2015) Sound-symbolism is disrupted in dyslexia: Implications for the role of cross-modal abstraction processes. In Noelle, D.; Dale, R.; Warlaumont, D.S. (ed.), *Proceedings of the 37th Annual Meeting of the Cognitive Science Society* (CogSci 2015), 602-607.

Fort, M., Lammertink, I., Peperkamp, S., Guevara-Rukoz, A., Fikkert, P., & Tsuji, S. (2018) A meta-analysis on the emergence of sound symbolism in early language acquisition. *Developmental Science*, *21*(5), e12659.

Fort, M., Martin, A., & Peperkamp, S. (2015) Consonants are more important than vowels in the bouba-kiki effect. *Language and Speech*, *58*(2), 247-266.

Fort, M., Weiß, A., Martin, A., & Peperkamp, S. (2013) Looking for the bouba-kiki effect in prelexical infants. In S. Ouni, F. Berthomier, & A. Jesse (Eds.), *The 12th International Conference on Auditory-Visual Speech Processing,* 71-76. INRIA.

Hayes, B., & White, J. (2013) Phonological naturalness and phonotactic learning. *Linguistic inquiry*, *44*(1), 45-75.

Hayes, B., & Wilson, C. (2008) A maximum entropy model of phonotactics and phonotactic learning. *Linguistic inquiry*, *39*(3), 379-440.

Imai, M., & Kita, S. (2014) The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1651), 1-13.

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), 1-11.

Köhler, W. (1929/1947) *Gestalt psychology*. New York, NY: Liveright Publishing.

Lockwood, G., & Dingemanse, M. (2015) Iconicity in the lab: A review of behavioral, developmental, and neuroimaging research into sound-symbolism. *Frontiers in Psychology*, *6*, 1-14.
Margiotoudi, K., Allritz, M., Bohn, M., & Pulvermüller, F. (2019) Sound symbolic congruency detection in humans but not in great apes. *Scientific Reports*, 9(1), 1-12.
Mateus, M. H. M., Falé, I., & Freitas, M. J. (2016) *Fonética e fonologia do Português, 2ª* edição. Lisboa: Universidade Aberta.
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.
Monaghan, P., & Fletcher, M. (2019) Do sound symbolism effects for written words relate to individual phonemes or to phoneme features?. *Language and Cognition*, *11*(2), 235-255.
Monaghan, P., Shillcock, R. C., Christiansen, M. H., & Kirby, S. (2014) How arbitrary is language?. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*, 20130299.
Nielsen, A. K., & Dingemanse, M. (2020) Iconicity in Word Learning and Beyond: A Critical Review. *Language and Speech*. https://doi.org/10.1177/0023830920914339.
Oberman, L. M., & Ramachandran, V. S. (2008) Preliminary evidence for deficits in multisensory integration in autism spectrum disorders: The mirror neuron hypothesis. *Social Neuroscience*, *3*(3-4), 348-355.
Occelli, V., Esposito, G., Venuti, P., Arduino, G. M., & Zampini, M. (2013) The Takete—Maluma phenomenon in autism spectrum disorders. *Perception*, *42*(2), 233-241.
Ozturk, O., Krehm, M., & Vouloumanos, A. (2013) Sound symbolism in infancy: Evidence for sound–shape cross-modal correspondences in 4-month-olds. *Journal of Experimental Child Psychology*, *114*(2), 173-186.
Pejovic, J., & Molnar, M. (2017) The development of spontaneous sound-shape matching in monolingual and bilingual infants during the first year. *Developmental Psychology*, *53*(3), 581-586.
Prickett, B. (2018) Complexity and Naturalness Biases in Phonotactics: Hayes & White (2013) Revisited. In G. Gallagher, M. Gouskova, & S. Yin (Eds.), *Supplemental Proceedings of the 2017 Annual Meeting on Phonology*. Washington, D.C.: Linguistic Society of America.
Ramachandran, V. S., & Hubbard, E. M. (2001) Synaesthesia--a window into perception, thought and language. *Journal of Consciousness Studies*, *8*(12), 3-34.
Rogers, S. K., & Ross, A. S. (1975) A cross-cultural test of the Maluma-Takete phenomenon. *Perception*, *4*(1), 105–106.
Saussure, F. D. (1959) *Course in general linguistics*. New York: Philosophical Library.
Selkirk, E. (1984) *Phonology and syntax: The relation between sound and structure*. Cambridge: The MIT Press.
Sidhu, D. M., & Pexman, P. M. (2018) Five mechanisms of sound symbolic association. *Psychonomic bulletin & review*, *25*(5), 1619-1643.
Silva, D. M. R., & Bellini-Leite, S. C. (2019) Cross-modal correspondences in sine wave: Speech versus non-speech modes. *Attention, Perception, & Psychophysics*. https://doi.org/10.3758/s13414-019-01835-z.
Styles, S. J., & Gawne, L. (2017) When does maluma/takete fail? Two key failures and a meta-analysis suggest that phonology and phonotactics matter. *i-Perception*, *8*(4), 1-17.
Vigário, M., & Falé, I. (1994) A Sílaba no Português Fundamental: Uma descrição e algumas considerações de ordem teórica. In *Actas do IX Encontro da Associação Portuguesa de Linguística* (pp. 465-477). Lisboa: APL/Colibri.
Westbury, C., Hollis, G., Sidhu, D. M., & Pexman, P. M. (2018) Weighing up the evidence for sound symbolism: Distributional properties predict cue strength. *Journal of Memory and Language*, *99*, 122-150.