

Global shape perception contributes to crossmodal correspondences

Yi-Chuan Chen^a, Pi-Chun Huang^b, & Charles Spence^c

^a Department of Medicine, Mackay Medical College, No.46, Sec.3, Zhongzheng Rd. Sanzhi Dist., New Taipei City, Taiwan, 252

^b Department of Psychology, National Cheng Kung University, 1 University Road, Tainan, Taiwan, 701

^c Department of Experimental Psychology, Anna Watts Building, University of Oxford Woodstock Road, Oxford, OX2 6GG, UK

Resubmitted to: *Journal of Experimental Psychology: Human Perception and Performance*

(Report)

Word count: 11,479 words (All included)

Date: 2020/10/11

Correspondence to: Dr. Pi-Chun Huang, Department of Psychology, National Cheng Kung University, No.1, University Road, Tainan City 701, Taiwan

Tel: +886-6-2757575 #56516; E-mail: pichun_huang@mail.ncku.edu.tw

Abstract

The Bouba/Kiki effect constitutes a classic sound-shape correspondence, with the meaningless sounds “Bouba” and “Kiki” being mapped onto smooth and spiky patterns, respectively. While it is commonly believed that the Bouba/Kiki effect is driven by the local rounded and angular features of a pattern, here we investigated the importance of an alternative level of visual processing—namely the global contours. We adopted compound radial frequency (RF) patterns and segmented them into convexities (outward curves) or concavities (inward curves). Note that convexities are more informative in terms of inferring the global contour than concavities. When the perceptual grouping of segments was facilitated by increasing their length, the grouping of convexities was more efficient than that of concavities as manifested by the closer matching judgments to the compound RF patterns. When we interfered with the perceptual grouping of segments by rotating each segment by 180°, the matching consensus of convexities was higher when they were presented in the original than in the reversed orientation. Hence, the Bouba/Kiki effect was susceptible to the factors modulating the perceptual grouping process going from segments to the global contour, suggesting that the Bouba/Kiki effect may occur at the global level of shape perception. Sound-shape correspondences would therefore seem to be expressed at multiple levels of information processing, furthering our understanding of the development, underpinning neural mechanisms, and applications of crossmodal correspondences.

1 Key words: Audiovisual, Sound symbolism, Radial-frequency pattern, Perceptual grouping, Global
2 contour

3

4

5 **Public Significance Statement**

6 The human brain searches for the relations between, and regularities of, incoming sensory signals in
7 order to reconstruct and better understand the physical world. One classic example is a matching
8 between sounds and shapes, as illustrated by the classic Bouba/Kiki effect, first reported in 1929 by
9 Wolfgang Köhler. This is where the meaningless speech sounds (i.e., “Bouba” and “Kiki”) are
10 mapped onto smooth and spiky patterns, respectively. In contrast to the general agreement that the
11 Bouba/Kiki effect is driven by the local rounded and angular visual features, we demonstrate for the
12 first time that the global contour that emerges from the grouping of segments is critical to this
13 phenomenon. Our study highlights the level of information processing at which this sound-shape
14 correspondence occurs in humans. These findings provide an intriguing direction for future research
15 examining the developmental trajectory, underpinning neural mechanisms, and applications of such
16 sound-shape correspondences.

17

1. Introduction

In order to make sense of the noisy multisensory world in which we find ourselves, the inputs from different senses belonging to a common source need to be integrated in order to construct a unified representation. This process is known as the *multisensory binding problem* (e.g., Chen & Spence, 2017; Rohe et al., 2019; Shams & Beierholm, 2010; Spence & Frings, 2019). In order to solve this problem, our brains not only use stimulus spatiotemporal structure (Parise et al., 2012; Stein & Meredith, 1993; Welch & Warren, 1980), but also extract certain co-occurrence regularities and non-arbitrary relationships between stimuli such as crossmodal correspondences (Marks, 2013; Spence, 2011; Spence & Sathian, 2020; Walker, 2012). The majority of crossmodal correspondences reflect mappings between simple features either along polar dimensions (such as higher-pitched tone and smaller-sized objects; Parise & Spence, 2009; Walker et al., 2010) or between perceptual categories (such as reddish foods/colours and sweetness or spiciness; Spence et al., 2015; Woods & Spence, 2016). One unique exception is a sound-shape correspondence between complex stimuli (see Figure 1 for examples): People typically match visual shapes with smooth contour to nonsense spoken words like “Maluma” or “Bouba”, while matching others with a spiky contour are matched to “Takete” or “Kiki” (e.g., Köhler, 1910, 1929; Ramachandran & Hubbard, 2001). This phenomenon, hereafter called the Bouba/Kiki effect, is commonly agreed to reflect the local rounded and angular features in each figure driving this correspondence (e.g., Holland & Wertheimer, 1964). Nevertheless, given the fact that the Bouba/Kiki effect was first proposed as an example of subjective experience in Gestalt psychology (Köhler, 1929), we would like to examine

1 whether this phenomenon is simply a collections of crossmodal correspondences at the feature
2 level, or any emergent property at a higher level of information processing dominate the
3 Bouba/Kiki effect.

4



5

6 Figure 1. Examples of smooth and spiky figures that have been used in previous studies of the Bouba/Kiki
7 effect.

8

9 At least three possible accounts at different levels of information processing have been
10 proposed for the Bouba/Kiki effect: the linguistic, semantic, and perceptual hypotheses (e.g., Marks,
11 1996; Ramachandran et al., 2020; Spence, 2011). According to the linguistic hypothesis, the
12 Bouba/Kiki effect is a developing outcome of sensory-motor connections during speech learning.
13 Specifically, smooth and spiky visual shapes resemble lip movements when uttering the vowel /u/
14 and /i/, respectively (see Ramachandran & Hubbard, 2001). However, given that preverbal infants
15 are sensitive to the Bouba/Kiki effect (e.g., Asano et al., 2015; Ozturk et al., 2013), an account at
16 the linguistic level cannot provide a full explanation of the phenomenon. Instead, it is suggested that
17 such sound-shape correspondences may provide a basis for language acquisition early in life (e.g.,
18 Imai & Kita, 2014; Maurer et al., 2006; though see Pejovic & Molnar, 2017).

1 The semantic hypothesis suggests that particular speech sounds are associated with certain
2 meaning or concepts, also known as “sound symbolism” or “perceptual metaphor” (Marks, 1996;
3 Nuckolls, 1999; Sidhu & Pexman, 2018). A classic example comes from Berlin’s (1994) study
4 showing that English speakers are able to categorize the words in an unknown Peruvian language
5 (Huambisa) into to “bird” or “fish” category at a level that is above chance. One of the possible
6 clues is that the first vowel of bird’s names often consisted of the vowel /i/ that is plausibly
7 associated with the ideas of “quick and rapid”; on the other hand, the names of fish often consisted
8 of the vowel /a/, plausibly associated with the ideas of “slow and smooth”. More recently, another
9 study demonstrated that English speakers have no problem consensually matching certain nonsense
10 words (e.g., “axittic”) to either one side of contrasting concepts (such as “sharp-round”, “large-
11 small”, and “masculine-feminine”, Westbury et al., 2018). Taken together, the above evidence
12 suggests that the concepts underpinning sound symbolism may be abstract, universal, and amodal
13 (Bremner et al., 2013; Davis, 1961; Marks, 1996). However, substantial differences of the
14 Bouba/Kiki effect have also sometimes been demonstrated in different cultures or countries (e.g.,
15 Chen et al., 2016; Rogers & Ross, 1975; Shang & Styles, 2017). In addition, reports from the early
16 blind suggest that they show less consensual matching between the sounds of “Bouba” and “Kiki”
17 to tactile smooth and spiky shapes than do sighted controls (Bottini et al., 2019; Fryer et al., 2014;
18 Sourav et al., 2019). Hence, the concepts underpinning the Bouba/Kiki effect are not completely
19 universal and amodal; instead, they are shaped by a person’s prior perceptual experience.

1 The perceptual hypothesis of the Bouba/Kiki effect aims to find out the relationships between
2 the perceptual features of sounds and shapes that drive the participant’s matching judgments. For
3 example, the power-spectrum analysis of speech sounds demonstrates that the spoken word “Kiki”,
4 as compared to “Bouba”, consists of stronger high-frequency signals (above 10,000 Hz, see Chen et
5 al., 2016). Hence, the Bouba/Kiki effect may be a form of the crossmodal correspondence between
6 lower-pitched sounds and rounded shapes whilst higher-pitched sounds and more angular shapes
7 (Hamilton-Fletcher et al., 2018; Marks, 1987). Such correspondences between pitch and shape have
8 been demonstrated in 4-month-old infants (Walker et al., 2010, 2014; though see Lewkowicz &
9 Minar, 2014), consistent with the emergence of the Bouba/Kiki effect in early infancy (Ozturk et
10 al., 2013). Taken together, the Bouba/Kiki effect, though partly underpinned by linguistic and
11 semantic mechanisms, may therefore originate from the perceptual stage of information processing.

12 The Bouba/Kiki effect has long been studied using hand-drawing patterns and a few of their
13 variations (e.g., Bremner et al., 2013; Maurer et al., 2006; see Nielsen & Rendall, 2013, for a
14 review). That is, thus far, very few studies have systematically manipulated the visual features in
15 this phenomenon (e.g., Sievers et al., 2019) compared to the accumulating studies that have
16 investigated the critical phonological and acoustic features driving the Bouba/Kiki effect (e.g.,
17 Bottini et al., 2019; Knoeferle et al., 2017; Nielsen & Rendall, 2011, 2013; Shang & Styles, 2017;
18 Silva & Bellini-Leite, 2019; Spector & Maurer, 2013; Westbury et al., 2018). To investigate this
19 issue, Chen et al. (2016) were the first to use radical frequency (RF) patterns (Wilkinson et al.,
20 1998), because their features can be systematically manipulated (see also Chen, et al., 2019). RF

1 patterns are closed contours with sinusoidal modulations along the circumference of a circle that are
2 frequently used to investigate shape perception (Wilkinson et al., 1998; see Figure 2A). It is
3 suggested that combining multiple RF patterns can generate a complex shape that resembles the
4 outline of a real object (Wilson & Wilkinson, 2002, Wilson et al., 2002; though see Schmidtman
5 & Freund, 2019).

6 Chen et al. (2016) manipulated the number and magnitude of sinusoidal modulations per circle,
7 as well as the spikiness in terms of adding harmonic triangular waveforms on top of each sinusoidal
8 modulation, in a stepwise manner. The results demonstrate that the participants were more likely to
9 match an RF pattern to “Kiki” rather than to “Bouba” when the frequency of RF patterns increased;
10 more specifically, the frequency of five was the boundary roughly separating the “Bouba” and
11 “Kiki” responses. Interestingly, it has been suggested that RF patterns are processed holistically
12 when the frequency is five or lower; otherwise, they are processed analytically by processing each
13 lobe individually and then pooling the information into a global pattern based on *probability*
14 *summation rule* (Hess et al., 1999; Loffler et al., 2003, see also Baldwin et al., 2016; Schmidtman
15 et al., 2019; Schmidtman & Kingdom, 2017 for alternative explanation for the processing of RF
16 patterns, e.g. pooling with additive summation rules or in the absence of pooling process). This
17 observation leads to an interesting question pertaining to the level of visual processing at which the
18 Bouba/Kiki effect occurs: Conventionally, researchers focused at the rounded or angular features of
19 a pattern; alternatively, the shape perception constructed by pooling these local features into a
20 global pattern may be what actually drives these sound-shape correspondences.

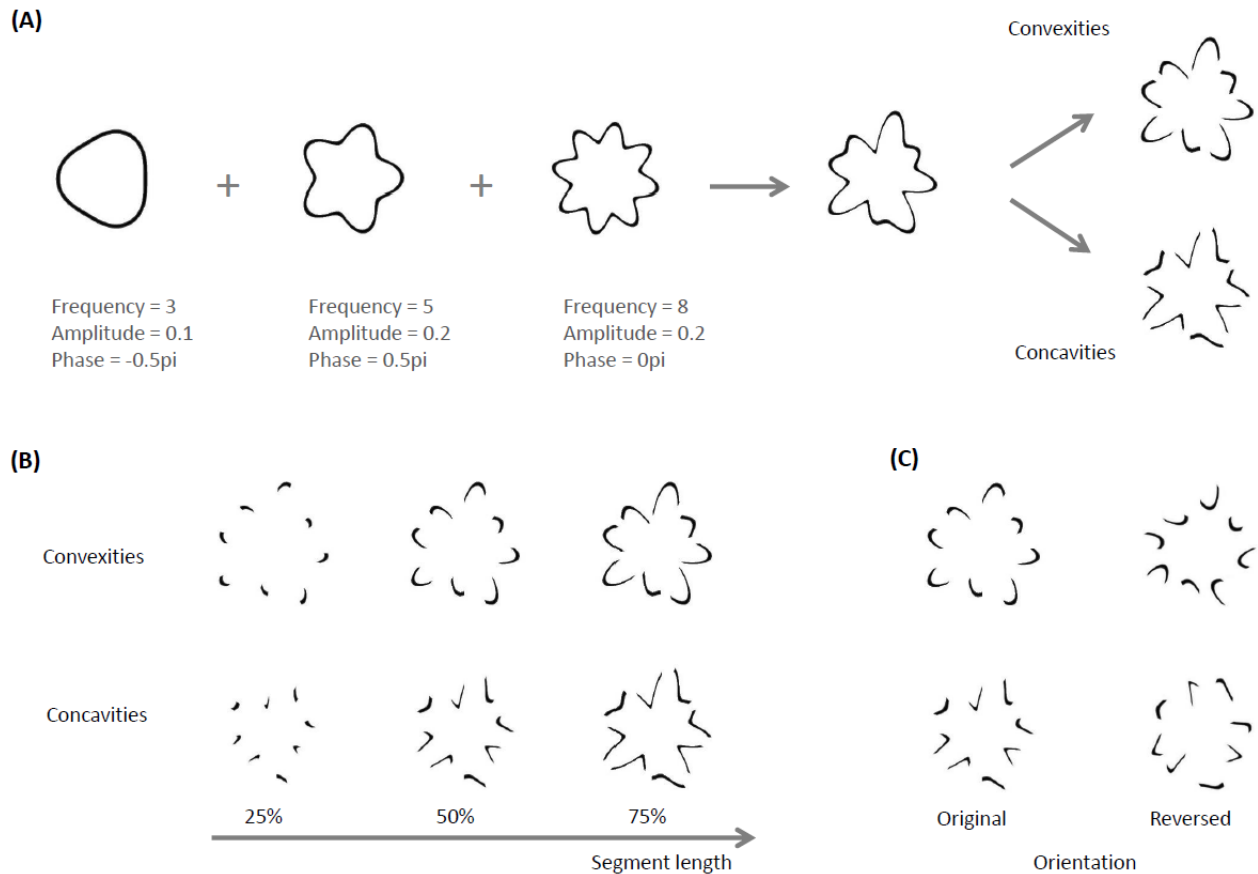


Figure 2. The visual stimuli used in the current study (A) Each compound RF pattern was created by combining three RF patterns using different amplitude and phase combinations (see Appendix). Then, the compound RF patterns was segmented into convexities (i.e., the outward curves) or concavities (i.e., the inward curves). (B) In Experiment 1, the segment length of convexities and concavities were manipulated. (C) In Experiment 2, the orientation of each segment was either in the original or reversed orientation.

To address this issue, we adapted variations of compound RF patterns used in Schmidtman et al. (2015): The compound RF patterns which combined three RF patterns additively gave rise to unfamiliar and asymmetrical patterns (see Figure 2A). Each compound RF pattern was then segmented into convexities (i.e., the outward curves) or concavities (i.e., the inward curves). The presented segments can be grouped and used to compute the unseen curvatures (e.g., Fantoni & Gerbino, 2003; Feldman, 1997). People find it easier to recognize a pattern when presented with its convexities rather than with concavities (Schmidtman et al., 2015, see Appendix for a replication;

1 though see Feldman & Singh, 2005). One reason for this is that convexities represent the maximum
2 numbers of curvatures and the outer boundary of a closed pattern. A visual pattern can be
3 reconstructed even when the end-points of convexities are linked using straight lines (Bell et al.,
4 2010; Schmidtman et al., 2015), following the minimal path of a gap (i.e., the *simplicity principle*
5 of perceptual grouping, Chater & Vitányi, 2003). An alternative possibility, based on amodal
6 completion, suggests that the imagery contour was more extrapolated (i.e., curvier) when the
7 occluded part was concave rather than when it was convex (Fantoni et al., 2005). Hence, when
8 grouping segments, interpolating the trajectories between convexities (i.e., imaging the missing
9 concavities) rather than between concavities (i.e., imaging the missing convexities) should generate
10 a contour that is more similar to the original pattern. This interpolation process aims to seek for an
11 optimal trajectory connecting nearby segments, consistent with the *likelihood principle* of
12 perceptual grouping (Feldman, 1999, 2009).

13 We therefore compared the participants' sound-shape matching judgments when viewing
14 segmented convexities, segmented concavities, or compound RF patterns. The hypothesis was that
15 if global contour is critical to Bouba/Kiki judgments, then the segments that are easily grouped
16 (such as convexities rather than concavities) would be matched to the same sound as the compound
17 RF patterns. In contrast, if local elements determine Bouba/Kiki judgments, then the presentation of
18 rounded or angular segments, rather than their grouping tendency, would seem to dominate
19 matching judgments instead.

2. Experiment 1

In Experiment 1, our aim was to examine the role of perceptual grouping from segments to the contour of a pattern in the Bouba/Kiki effect. Two factors that presumably influence the grouping process were manipulated: curvature type (convexities or concavities) and segment length. Specifically, convex rather than concave segments, and longer rather than shorter segments, should facilitate the construction of a pattern representation (see Schmidtman et al., 2015, and Appendix). Therefore, our assumption was that, if global contour rather than local features drive the Bouba/Kiki effect, the sound-shape matching judgments of segments that are easier to group into the whole contour would be closer to the matching judgment of compound RF pattern.

2.1 Methods

2.1.1 Participants

A group of 123 participants (mean age = 19.5 ± 1.5 years, range = 18-24 years, 72 females) took part in Experiment 1. All were students at National Cheng Kung University in Taiwan and received course credit in return for taking part. They were naïve as to the purpose of the study. The participants gave their informed consent prior to the start of the experiment. The procedures were carried out in accordance with the Declaration of Helsinki and were approved by the Department of Psychology, National Cheng Kung University.

An analysis using G*Power for chi-square test (version 3.1.9.4; Faul et al., 2007) suggests that the power reaches 0.85 when testing 100 participants with type-I error (α) setting at the 0.05 level

1 and a medium level of effect size (Cohen's $\omega = 0.3$). Hence, at least a hundred participants were
2 planned to test in each experiment (see Chen et al., 2019).

3 *2.1.2 Stimuli and Design*

4 The visual stimuli were compound RF patterns created by combining three RF patterns
5 (frequencies of three, five, and eight cycle/circle) using one of six amplitude combinations and one
6 of four phase combinations (see Figure 2A and Appendix for the procedure of creating these stimuli).
7 Hence, there were a total of 24 compound RF patterns. Each compound RF pattern was segmented
8 into eight convexities or concavities with equal angles along the circumference, and each segment
9 presented with 25%, 50%, 75%, or 100% of the length (see Figure 2B for examples). Taken together,
10 there are 168 RF patterns as visual stimuli, including 24 compound RF patterns (i.e., 100% segment
11 length) and 144 segmental RF patterns (6 amplitude combinations x 4 phase combinations x 2
12 convexities/concavities x 3 segment lengths).

13 The auditory stimuli consisted of the spoken nonsense words “Bouba” and “Kiki” produced by
14 a female native English speaker three times for each word with slightly different prosody (32 bit
15 mono; 44,100 Hz digitization). All six sound files were edited to the same length (400 ms) and their
16 sound pressure level (in terms of the value of root mean square) were equalized.

17 *2.1.3 Procedure*

18 Experiment 1 was conducted on an online platform via the Adobe Flash based Xperiment
19 software 2.0 (<https://www.xpt.mobi/>). The participants were requested to switch to full screen mode

1 and confirmed that they could hear the sounds clearly (by typing in three digits that they heard) before
2 the experiment started.

3 On each trial, a visual pattern was presented in the center of the monitor, and the spoken non-
4 word “Bouba” and “Kiki” were presented auditorily – their order of presentation was randomized on
5 a trial-by-trial basis. The participants had to judge whether the spoken words “Bouba” or “Kiki”
6 provided a better match for the seen pattern – they had to choose one or the other by clicking on the
7 button labeled “Sound 1” or “Sound 2” on the monitor in order to proceed. Each participant completed
8 the first block containing the 144 segmental RF patterns, followed by the second block containing the
9 24 compound RF patterns, and the final block containing the two figures commonly used to test the
10 Bouba/Kiki effect (see Figure 1; Bremner et al., 2013; Chen et al., 2016). The presentation order of
11 the patterns in each block was completely randomized.

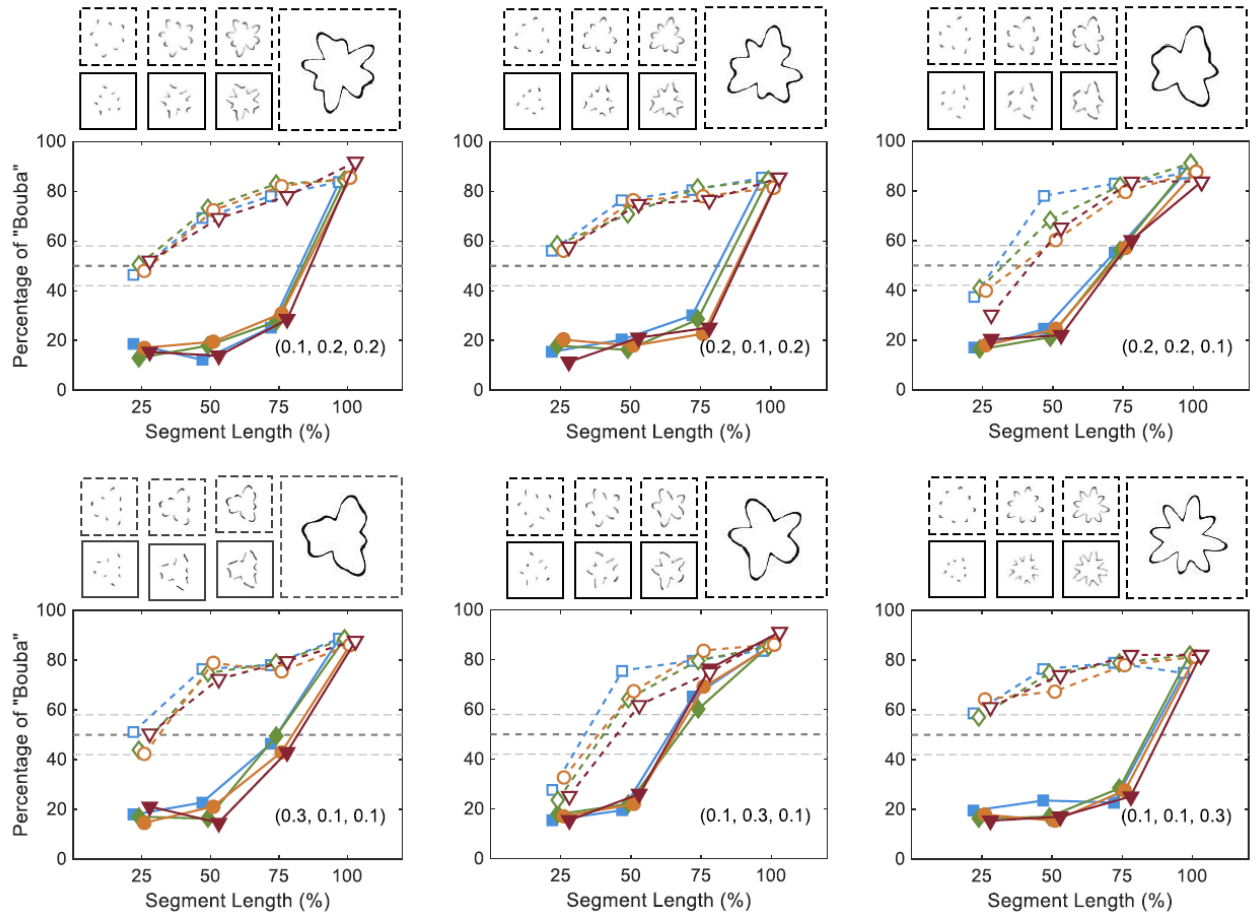
12 **2.2 Results**

13 We first confirmed that this group of participants demonstrated typical sound-shape
14 correspondences: When they were tested with the two commonly-used figures, 88.6% of the
15 participants matched the rounded figure to “Bouba” ($\chi^2(1) = 73.37, p < .001$, Cohen’s $\omega = 0.77$), and
16 74.0% of participants matched the spiky figure to “Kiki” ($\chi^2(1) = 28.30, p < .001$, Cohen’s $\omega = 0.48$).
17 Hence, this group of participants demonstrated the classic Bouba/Kiki effect.

18 The proportion answering “Bouba” for each RF pattern was plotted as a function of segment
19 length in Figure 3. The results of compound and segmental RF patterns belonging to the same
20 amplitude combination (including four phase combinations) were plotted in the same figure. The light

1 grey dashed lines indicate the boundaries of undetermined matching. Three main results can be
 2 summarized when inspecting Figure 3: First, all of the 24 compound RF patterns were matched to
 3 “Bouba” (ranging from 74.8% to 91.9%, all $X^2(1) > 30.25$, $p < .001$, Cohen’s $\omega > 0.49$). Second,
 4 presenting the convexities or concavities of the segmental RF patterns demonstrated different
 5 matching results (see Table 1). Specifically, among the 24 segmental RF patterns of 25% convexities,
 6 two were matched to “Bouba” (61.0% and 64.2%, both $X^2(1) > 5.92$, $p < .05$, Cohen’s $\omega > 0.21$), 14
 7 were undetermined (percentage of “Bouba” response ranging from 42.3% to 58.5%, all $X^2(1) < 3.59$,
 8 $p > .05$, Cohen’s $\omega < 0.17$), and the remaining eight were matched to “Kiki” (percentage of “Bouba”
 9 response ranging from 23.6% to 40.7%, all $X^2(1) > 4.30$, $p < .05$, Cohen’s $\omega > 0.18$). Hence, among
 10 the three types of matching results (“Bouba”, undetermined, or “Kiki”), the majority of these 24
 11 segmental RF patterns of 25% convexities were undetermined ($X^2(2) = 9.00$, $p < .05$, Cohen’s $\omega =$
 12 0.61). When 50% and 75% convexities were presented, all of the segmental RF patterns were matched
 13 to “Bouba” (ranging from 60.2% to 83.7%, all $X^2(1) > 5.08$, $p < .05$, Cohen’s $\omega > 0.20$). In contrast,
 14 when 25% and 50% concavities were presented, all of these segmental RF patterns were matched to
 15 “Kiki” (percentage of “Bouba” response ranging from 11.4% to 26.0%, all $X^2(1) > 28.30$, $p < .001$,
 16 Cohen’s $\omega > 0.47$). Among the 24 segmental RF patterns of 75% concavities, five were matched to
 17 “Bouba” (ranging from 60.2% to 76.4%, all $X^2(1) > 5.08$, $p < .05$, Cohen’s $\omega > 0.20$), seven were
 18 undetermined (percentage of “Bouba” response ranging from 43.1% to 56.9%, all $X^2(1) < 2.35$, $p >$
 19 .12, Cohen’s $\omega < 0.14$), and the remaining 12 were matched to “Kiki” (percentage of “Bouba”

1 response ranging from 22.8% to 30.9%, all $X^2(1) > 17.95$, $p < .001$, Cohen's $\omega > 0.38$). None of the
 2 three matching types was dominant ($X^2(2) = 3.25$, $p = .20$, Cohen's $\omega = 0.37$).



4 **Figure 3.** Mean percentage of “Bouba” responses of each compound and segmental RF patterns in Experiment
 5 1. The open shapes and dashed lines represent the results for convexities, and the filled shapes and solid lines
 6 represent the results for concavities. Each shape-color combination represents one of the phase combinations.

Table 1. The results of Experiment 1. The number of patterns (out of 24) that were matched to “Bouba”, “Kiki”, or undetermined in each of the curvature x segment length conditions, and the Chi-square test of the distribution violating equal probability for each of the three response types. *: the majority response type in each condition.

Curvature	Segment length	Matching responses			$X^2(2)$	p	Cohen’s ω
		Bouba	Undetermined	Kiki			
Convexities	25%	2	14*	8	9.00	< .05	0.61
	50%	24*	0	0	48.00	< .001	1.41
	75%	24*	0	0	48.00	< .001	1.41
Concavities	25%	0	0	24*	48.00	< .001	1.41
	50%	0	0	24*	48.00	< .001	1.41
	75%	5	7	12	3.25	= .20	0.37

Third, when the segment length increased, it seems that the increasing rate of the proportion of “Bouba” responses (i.e., the majority response of the compound RF patterns) differed when either convexities or concavities were presented. In order to examine this, logit regressions were used to predict the participant’s matching judgments by using the curvatures (convexities or concavities) and segment length (25%, 50%, and 75%) as the predictors, while the random variables were participants and the phase combinations nested in amplitude combinations. The linear mixed-effect model (lme4 package, version 1.1-8, Bates et al., 2015) in R (version 3.2.1) was used to fit the data using the maximum likelihood method. It was also used to apply parametric bootstrapping method 1,000,000 times in order to derive the standard error (SE) for each coefficient. Two models were compared (see Table 2): Model 1 assumed the same coefficient of segment length for convexities and concavities; Model 2 assumed different coefficients of segment length for convexities (the Segment Length) and

1 concavities (the Segment Length + Δ Segment Length). Both Models 1 and 2 demonstrated significant
2 positive coefficients for curvature and segment length, suggesting that both factors effectively
3 modulated people's matching judgments. Comparing the two models, Model 2 provided a better fit,
4 suggesting that the coefficient of segment length was higher for convexities than concavities a (given
5 that the coefficient of Δ Segment Length was negative). In summary, the presentation of convexities
6 rather than concavities, and the presentation of longer segments, led to a higher probability of "Bouba"
7 responses; furthermore, the increasing segment length of convexities led to a more rapid increment
8 of "Bouba" responses than concavities.

9

10 Table 2. The model comparison in Experiment 1. The coefficient, 95% confidence interval (CI, in brackets), degrees of
11 freedom (*df*), and goodness of fit of Models 1 and 2, and their comparison using Chi-square test. **: $p < .01$; ***: $p <$
12 .001.

Models	Curvature	Segment length	Δ Segment length	Constant	<i>df</i>	Log likelihood	Deviance	$X^2(1)$	<i>p</i>
1	1.96*** [1.89, 2.03]	3.12*** [2.94, 3.30]	-	-4.72*** [-4.92, -4.52]	3	-9794.4	19588.7	8.4	< .005
2	1.69*** [1.50, 1.89]	3.36*** [3.12, 3.61]	-0.52** [-0.87, -0.17]	-4.31*** [-4.65, -3.96]	4	-9790.2	19580.3		

13 Model 1: Response ~ 1 + Curvature + Segment length + (1|Subject) + (1 + 3RF + 5RF + 8RF|Phase)

14 Model 2: Response ~ 1 + Curvature + Segment length + Δ Segment Length + (1|Subject) + (1 + 3RF + 5RF + 8RF|Phase)

15

1 **2.3 Discussion**

2 In Experiment 1, we examined whether the grouping process of segments modulates the sound-
3 shape matchings in the Bouba/Kiki effect. We used the compound RF patterns which were
4 predominantly matched to “Bouba”. When these compound RF patterns were presented segmentally
5 with only convexities or concavities, their convexities were more likely to be matched to “Bouba”;
6 specifically, all of the segmental RF patterns with 50% or 75% convexities were reliably matched to
7 “Bouba”, whereas the segmental RF patterns with 75% concavities did not show a dominant matching
8 result. Finally, the proportion of “Bouba” responses increased with increasing segment length more
9 rapidly when convexities rather than concavities were presented.

10 Our assumption here was that given the compound RF patterns predominantly matched to
11 “Bouba”, the segments that were easier to reconstruct the compound RF patterns would be more
12 likely to be matched to “Bouba” as well. Consistent with this hypothesis, we demonstrated that
13 convexities rather than concavities and longer segments that are easier to group into a global contour
14 were more likely to be matched to “Bouba”. Furthermore, these two factors interact: When increasing
15 the segmental information of a contour (i.e., increasing the segment length), the information of
16 convexities was pooled together more efficiently than concavities. This result is therefore consistent
17 with Schmidtmann et al. (2015) who demonstrate that the convexities, rather than the concavities,
18 constitute the critical information underpinning shape perception. We further extended their finding
19 in that the grouping process was more efficient for convexities than for concavities when constructing
20 a contour as manifested by the proportion of “Bouba” judgments. The advantage of pooling

1 information for convexities over concavities when reconstructing the complete pattern can be
2 explained by interpolation mechanisms proposed by Schmidtmann et al. (2015) and Fantoni et al.
3 (2005). Specifically, either using straight lines or inferencing optimal trajectory to link the endpoints
4 of segments, convexities compared to concavities would create a pattern that is more similar to the
5 original one.

6 The grouping of segments into contour therefore seems to be a critical process in the Bouba/Kiki
7 effect, especially for those patterns being matched to “Bouba”. That said, before concluding that the
8 global contour rather than local features of a pattern drive the Bouba/Kiki effect, one alternative
9 explanation needs to be excluded: The convexities are the outward curves while the concavities, the
10 inward curves, and the former are smoother by nature in the compound RF patterns used in the present
11 study. Hence, the possibility that the local rounded and angular curvatures (i.e., convexities and
12 concavities, respectively) drive the Bouba/Kiki effect remains. We excluded this alternative
13 explanation in Experiment 2.

14

15 **3. Experiment 2**

16 In Experiment 2, we aimed to further examine whether the higher probability of “Bouba”
17 responses for convexities than for concavities was genuinely attributable to the perception of global
18 contour, rather than the local features. Specifically, the assumption put forward here is that it would
19 be easier for participant to make a “Bouba” response when viewing the convexities of a compound
20 RF patterns in the original orientation; by contrast, these convexities would no longer be

1 predominantly matched to “Bouba” when the grouping of these convexities is interrupted. The
2 concavities, however, given the fact that they are hard to reconstruct the global contour, they would
3 be mainly matched to “Kiki” in both original and reversed orientations.

4 In contrast to Experiment 1 which was conducted online, Experiment 2 was conducted in the
5 laboratory. Hence, the physical properties of the stimuli (including the monitor resolution, the size
6 of each visual pattern, and the volume of speech sound, etc.) were all well-controlled between
7 participants.

8 **3.1 Methods**

9 A new group of 101 participants (mean age = 20.02 ± 1.19 years, range = 18-23 years, 46
10 females) from the same participant pool took part in Experiment 2. Twelve compound RF patterns
11 were created by combining three, five, and eight cycle/circle with the 0.1, 0.2 and 0.2 amplitude
12 combinations and four phase combinations (the patterns in Figure 3 upper panel). In order to
13 increase the generalization of our results, each compound RF pattern was segmented into five or
14 eight convexities or concavities with equal angle along the circumference. Each segment was
15 presented with 50% of the length. The reversed pattern of each convexity and concavity was created
16 by rotating the local segment for 180 degrees along the radial axis (see Figure 2C for the examples).
17 In this case, their local features remained the same while the reversed segments were not possible to
18 be grouped and reconstruct the original compound RF pattern. Taken together, there were a total of
19 96 segmental RF patterns (12 compound RF patterns x 2 five/eight segments x 2
20 convexities/concavities x 2 original/reversed).

1 The stimulus was presented on a 24-inch LCD monitor (1280 x 720 pixels) and controlled by a
2 PC compatible with Psychophysics Toolbox Version 3 (Brainard, 1997) under the MATLAB
3 environment (The Mathworks, Matric, MA, USA). With viewing distance of 60 cm, the size of the
4 RF pattern was 15 degrees. The auditory stimulus was presented at a comfortable sound level by a
5 pair of speakers that was placed 5 cm left and right from the monitor. On each trial, the fixation
6 point was presented for 100 ms and then a blank for 100 ms, followed by the visual pattern and the
7 sound. The presentation order of the sounds was randomized on a trial-by-trial basis. The visual
8 pattern disappeared after the participant responded. Each participant completed the first block
9 containing the 96 segmental RF patterns, then the second block containing the 12 compound RF
10 patterns, and the final block containing the two figures commonly used to test the Bouba/Kiki effect
11 (Bremner et al., 2013; Chen et al., 2016). The presentation order of the patterns in each block was
12 completely randomized.

13 **3.2 Results**

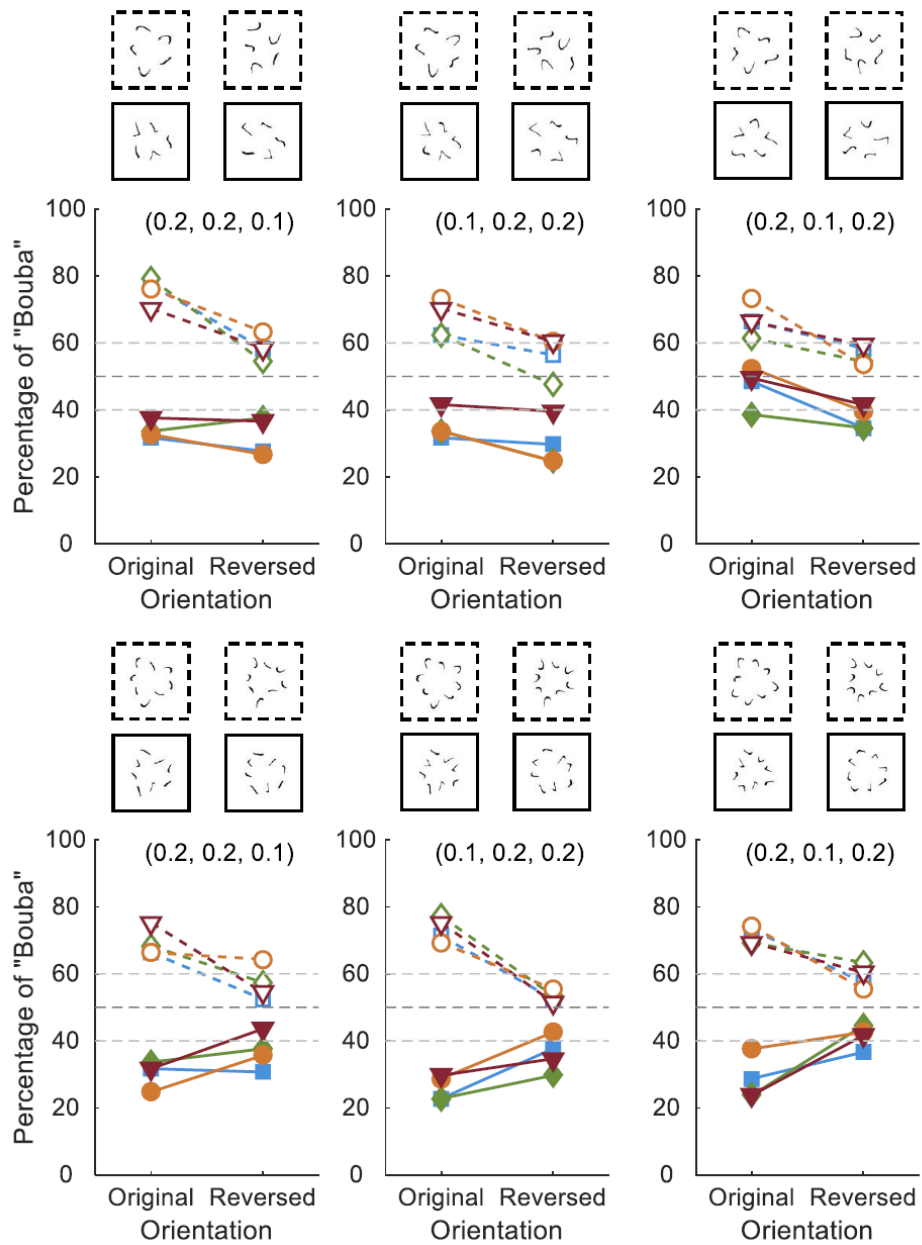
14 When the participants in Experiment 2 were tested using the two commonly-used figures,
15 54.5% of participants matched the rounded figure to “Bouba” ($\chi^2(1) = 0.80, p = .37$, Cohen’s $\omega =$
16 0.09), and 82.2% of participants matched the spiky figure to “Kiki” ($\chi^2(1) = 41.83, p < .001$,
17 Cohen’s $\omega = 0.64$). These participants therefore demonstrated a weak Bouba/Kiki association
18 mainly resulting from the consensual matching between the spiky pattern and “Kiki”.

19 Among the 12 compound RF patterns, six were predominantly matched to “Bouba” (ranging
20 from 61.4% to 77.2%, all $\chi^2(1) > 5.23, p < .05$, Cohen’s $\omega > 0.22$). Five were mainly matched to

1 “Bouba”, though this trend was not statistically significant (ranging from 53.5% to 59.4%, $X^2(1) <$
 2 $3.57, p \geq .059$, Cohen’s $\omega < 0.19$). There was no difference in the extent to which the last one was
 3 matched (the percentage of “Bouba” response was 49.5%, $X^2(1) = 0.01, p = .92$, Cohen’s $\omega = 0.01$).
 4 Figure 4 presents the proportion of “Bouba” responses as the Y-axis and the conditions of
 5 original or reversed segments as the X-axis. The results of segmental RF patterns with the same
 6 amplitude combination (including four phase combinations) is plotted in the same figure. When the
 7 segmental RF patterns were original convexities, all of the 12 patterns were matched to “Bouba”
 8 (ranging from 61.4% to 79.2%, all $X^2(1) > 5.23, p < .05$, Cohen’s $\omega > 0.22$) in both 5- and 8-
 9 segment condition. In contrast, when the convexities were reversed, only three segmental RF
 10 patterns were matched to “Bouba” (ranging from 60.4% to 64.4%, all $X^2(1) > 4.36, p < .05$, Cohen’s
 11 $\omega > 0.20$), while the majority (i.e., the remaining nine patterns) were undetermined (ranging from
 12 47.5% to 59.4%, all $X^2(1) < 3.57, p \geq .059$, Cohen’s $\omega < 0.19$), in both 5- and 8-segment condition.
 13 Hence, when the convexities were rotated from original to reversed orientation, the majority of the
 14 segmental RF patterns were less likely to be matched to “Bouba” (see Table 3). When the
 15 segmental patterns were original concavities, most of the patterns were matched to “Kiki” (8 and 12
 16 in the 5- and 8-segment condition, respectively, the percentage of “Bouba” response ranging from
 17 22.8% to 38.6%, all $X^2(1) > 5.23, p < .05$, Cohen’s $\omega > 0.22$), while the remaining were
 18 undetermined (the percentage of “Bouba” response ranging from 41.6% to 52.5%, all $X^2(1) < 2.86,$
 19 $p > .09$, Cohen’s $\omega < 0.17$). Similarly, when the concavities were reversed, most of the patterns
 20 were matched to “Kiki” as well (11 and 7 in the 5- and 8-segment condition, respectively, the

1 percentage of “Bouba” response ranging from 24.8% to 39.6%, all $X^2(1) > 4.36$, $p < .05$, Cohen’s ω
2 > 0.20), while the remaining were undetermined (the percentage of “Bouba” response ranging from
3 41.6% to 44.6%, all $X^2(1) < 2.86$, $p > .09$, Cohen’s $\omega < 0.17$). Hence, when the concavities were
4 presented either in the original or reversed orientation, the majority of segmental RF patterns were
5 matched to “Kiki” (see Table 3).

6



7

Figure 4. Mean percentage of “Bouba” responses of each segmental RF patterns in Experiment 2. There were five segments was in the upper panel, and eight in the lower panel. The open shapes and dashed lines represent the results for convexities, and the filled shapes and solid lines represent the results for concavities. Each shape-color combination represents one of the phase combinations.

Table 3. The results of Experiment 2. The number of patterns (out of 12) that were matched to “Bouba”, “Kiki”, or undetermined in each of the Number of segments x Curvature x Orientation conditions, and the Chi-square test of the distribution violating equal probability for each of the three response types. *: the majority response type in each condition.

Number of segments	Curvature	Orientation	Matching responses			$X^2(2)$	p	Cohen's ω
			Bouba	Undetermined	Kiki			
5	Convexities	Original	12*	0	0	24.00	< .001	1.41
		Reversed	3	9*	0	10.50	< .01	0.94
	Concavities	Original	0	4	8*	8.00	< .05	0.82
		Reversed	0	1	11*	18.50	< .001	1.24
8	Convexities	Original	12*	0	0	24.00	< .001	1.41
		Reversed	3	9*	0	10.50	< .01	0.94
	Concavities	Original	0	0	12*	24.00	< .001	1.41
		Reversed	0	5	7*	6.50	< .05	0.74

We then further verified the distinct influences of reversing segments in convexities and concavities using logit regressions. The three predictors were number of segment (5 or 8), curvature (convexities or concavities), and orientation (original or reversed), while the random variables were participants and phase combinations nested in amplitude combinations. Similar to Experiment 1, two models were compared (see Table 4): Model 1 assumed that the same coefficient of orientation

1 for the convexities and concavities; in contrast, Model 2 assumed that different coefficients of
2 orientation in the convexities (the Orientation) and concavities (the Orientation + Δ Orientation).
3 The results of the model comparison revealed that Model 2 provides a better-fit, suggesting that
4 reversing the segments induced different effects on matching judgments for the convexities and
5 concavities. Given that the values of coefficients of Orientation and Δ Orientation factors were
6 opposite values, we then conducted two linear mixed-effect models using number of segment and
7 orientation factors as predictors for convexities and concavities, separately (see Table 5). The
8 coefficient of orientation was negative for the convexities ($-0.70, p < .001$), thus suggesting that
9 reversing each convex segment significantly reduced the probability of “Bouba” responses.
10 However, the coefficient of Orientation was not significant for the concavities ($0.11, p = .10$),
11 suggesting that reversing each concave segment did not influence the matching response (i.e.,
12 mainly to “Kiki” in both orientations). Nevertheless, the coefficient of the number of segments
13 happened to be significant for the concavities ($-0.05, p < .05$), thus suggesting that when the
14 compound RF patterns were cut into more concave segments, they would be more likely to be
15 matched to “Kiki”.

1 Table 4. The model comparison in Experiment 2. The coefficient, 95% confidence interval (CI, in brackets), degrees of freedom (*df*), and goodness of fit of Models 1 and 2, and their
 2 comparison using Chi-square test. **: $p < .01$; ***: $p < .001$.

Model	Number of segment	Curvature	Orientation	Δ Orientation	Constant	<i>df</i>	Log likelihood	Deviance	$X^2(1)$	<i>p</i>
1	-0.02	1.24***	-0.26***	-	-0.35**	4	-6224.7	12449.4	68	< .001
	[-0.05, 0.01]	[1.15, 1.32]	[-0.35, -0.18]		[-0.58, -0.13]					
2	-0.02	1.60***	-0.62***	0.71***	-0.53***	5	-6190.7	12381.4		
	[-0.05, 0.01]	[1.48, 1.73]	[-0.74, -0.50]	[0.54, 0.88]	[-0.76, -0.30]					

3 Model 1: Response ~ 1 + Number of segments + Curvature + Orientation + (1|ID) + (1 + 3RF + 5RF + 8RF|Phase)

4 Model 2: Response ~ 1 + Number of segments + Curvature + Orientation + Δ Orientation + (1|ID) + (1 + 3RF + 5RF + 8RF|Phase)

5

6 Table 5. Predicting the participants' "Bouba" responses using Number of segment and Orientation as
 7 predictors for convexities and concavities separately in Experiment 2. The coefficient, 95% confidence
 8 interval (CI, in brackets). *: $p < .05$; ***: $p < .001$.

	Curvature	Number of segments	Orientation	Constant
Convexities		0.01	-0.70***	1.04***
		[-0.04, 0.05]	[-0.83, -0.57]	[0.68, 1.40]
Concavities		-0.05*	0.11	-0.40*
		[-0.09, -0.004]	[-0.02, 0.23]	[-0.77, -0.03]

9 Model: Response ~ 1 + Number of segments + Orientation + (1|ID) + (1 + 3RF + 5RF + 8RF|Phase)

10

1 When comparing the matching judgments of the patterns used in both Experiments 1 and 2
2 (The rounded shape in Figure 1, and the 12 compound RF patterns in Figure 3 upper panel), these
3 patterns were matched to “Bouba” with a higher degree of consensus in Experiment 1 than in
4 Experiment 2. We therefore examined the participants’ overall binomial distributions between the
5 two alternative responses in each experiment. The total numbers of “Bouba” and “Kiki” responses
6 in Experiment 1 were 10,806 (51.7%) and 10,104 (48.3%), respectively ($\chi^2(1) = 23.57, p < .001$,
7 Cohen’s $\omega = 0.03$), and those in Experiment 2 were 5,517 (49.7%) and 5,593 (50.3%), respectively
8 ($\chi^2(1) = 0.52, p = .47$, Cohen’s $\omega = 0.007$). Hence, the participants in Experiment 2 tended to
9 equalize their “Bouba” and “Kiki” responses, and this tendency may have reduced the proportion of
10 “Bouba” responses for certain patterns, especially in the later two blocks tested with compound RF
11 patterns and commonly-used figures. That said, this tendency to balance “Bouba” and “Kiki”
12 responses should not influence our main conclusion that reversing each segment significantly
13 reduced the “Bouba” responses of convexities rather than concavities given that these segmental RF
14 patterns were tested within the first block in a random order.

15 **3.3 Discussion**

16 The goal in Experiment 2 was to examine whether interrupting the grouping process of
17 segmental RF patterns would influence participants’ sound-shape matching judgments. To do so,
18 the convexities or concavities were presented in either the original orientation (i.e., they can be
19 grouped into the compound RF pattern), or else in the reversed orientation (i.e., the grouping
20 process is interrupted). The results demonstrated that the convexities in the original orientation were

1 predominantly matched to “Bouba”, while those in the reversed orientation became undetermined.
2 Hence, the proportion of “Bouba” responses was significantly reduced when the segment of the
3 convexities were reversed, suggesting that their grouping process is critical when making “Bouba”
4 judgments. In contrast, the concavities in either orientation were predominantly matched to “Kiki”,
5 and the matching responses were insensitive to reversing each segment. In addition, the proportion
6 of “Kiki” responses was increased when the number of segments increased. Combining these results
7 suggests whether and how segments were grouped was not essential when making “Kiki”
8 judgments.

9 The results of Experiment 2 therefore highlight a significant conclusion that, it was not the
10 rounded segments themselves (i.e., the convexities), but the smooth global contour constructed by
11 these segments gave rise to the matching judgment of “Bouba”. In contrast, the angular segments
12 (i.e., the concavities), no matter whether it was easy to construct a pattern with them or not,
13 remained predominantly matched to “Kiki”. Taken together, the pattern predominantly matched to
14 “Bouba” should satisfy two conditions: its local segments are rounded rather than angular, and
15 critically, these rounded segments can be integrated into a smooth global contour.

16

17

4. General Discussion

18 The present study examined the perceptual basis of the Bouba/Kiki effect. Specifically, as a
19 prototypical example of sound-shape correspondences, we examined whether the Bouba/Kiki effect
20 emerged at the level of local features or global contour in the pathway of visual information

1 processing. We used segmental RF patterns to examine this issue, which constitutes a novel set of
2 visual stimuli in the literature on the Bouba/Kiki effect. Experiment 1 was an online study that the
3 participants viewed a visual pattern and judged whether “Bouba” or “Kiki” that they heard provided
4 a better match to the pattern. For the compound RF patterns that were predominantly matched to
5 “Bouba”, the participants were more likely to make the same response when viewing their
6 convexities rather than concavities. In addition, the proportion of “Bouba” responses increased
7 more rapidly when increasing segment length of convexities than concavities. These results
8 therefore demonstrate that the sound-shape matching of “Bouba” judgments was sensitive to the
9 factors modulating the perceptual grouping process of the segments. Experiment 2 was laboratory-
10 based, and the stimuli presented to each participant were carefully controlled. The “Bouba”
11 response was significantly decreased when convexities were reversed in which case the grouping
12 process was interrupted; in contrast, concavities were predominantly matched to “Kiki” irrespective
13 of their orientations. These results again suggest that the grouping process studied here is critical to
14 the Bouba/Kiki effect, and especially to the “Bouba” judgments. Taken together, we demonstrated
15 that the grouping process of convexities is critical for a pattern to be matched to “Bouba”. We
16 therefore suggest that the perceptual mechanism underlying the Bouba/Kiki effect is a perceptual
17 grouping process from local features to global contour, rather than simply a stage representing local
18 rounded or angular features. Hence, the Bouba/Kiki effect emerges in the visual feedforward
19 processing involving shape perception, the so-called “mid-level representation” (Hubel & Wiesel,
20 1962; Riesenhuber & Poggio, 1999; Serre et al., 2007; Van Essen & Gallant, 1994).

1 In the present study, we mainly used the compound RF patterns matched to “Bouba” rather
2 than to “Kiki”; nevertheless, participants seemed to have no problem matching each pattern to
3 either “Bouba” or “Kiki” with similar probability over the whole testing session. These judgments
4 should not be random responses because we can predict participants’ matching judgments using
5 those factors that pertain to perceptual grouping (convexities/concavities and segment length in
6 Experiment 1, and convexities/concavities and segment orientation in Experiment 2). Interestingly,
7 while the segments which can be better grouped into a pattern were more likely to be matched to
8 “Bouba”, others being hard to group were predominantly matched to “Kiki” instead. These results
9 therefore question the relationship between the patterns matched to “Bouba” and “Kiki”,
10 respectively: Instead of being represented at the same level along a contrasting rounded-angular
11 dimension, they are plausibly represented hierarchically at more global and local levels. The latter
12 notion suggests that “Bouba” patterns are those with elements that are easily connected together and
13 then integrated into a smooth global contour, either following the simplicity principle (e.g.,
14 Schmidtman et al., 2015) or the likelihood principle (Fantoni et al., 2005). That is, “Bouba”
15 patterns are the “good form” or “Prägnanz” when using the Gestalt psychology term (Todorovic,
16 2008; Wertheimer, 1938). In contrast, other patterns with jagged, fragmentary elements that are not
17 easily grouped into a simple shape would be matched to “Kiki” instead. In recent years, the
18 simplicity and likelihood principles are suggested to be conjointly quantified using Bayesian
19 approach (Froyen et al., 2015; Wagemans et al., 2012). The segmental RF patterns perhaps provide
20 a useful set of stimuli with which to test a hypothesis that convexities rather than concavities should

1 lead to an optimal solution of boundary that is close to the original RF patterns efficiently. One
2 direction for future research is therefore to examine whether such a Bayesian approach can be
3 extended from visual to auditory perceptual grouping (e.g., Kubovy & Van Valkenburg, 2001), and
4 further to crossmodal perceptual grouping in a hierarchical framework (e.g., Froyen et al., 2015; see
5 Spence & Chen, 2012, for a discussion, and Rohe et al., 2019; Rohe & Noppeney, 2015, for
6 applications).

7 The hierarchical notion of “Bouba” and “Kiki” is perhaps useful to explain an interesting
8 observation in the Bouba/Kiki literature: In most studies, participants have been presented with both
9 rounded and angular patterns side-by-side, as well as typically hearing both sounds in each trial, and
10 they were asked for the better matching ways between the sounds and shapes (e.g., Bremner et al.,
11 2013; Holland & Wertheimer, 1964; Maurer et al., 2006). Hence, it is unclear whether both pairings
12 were matched equally well, or whether instead one pair might make more sense than the other.
13 However, when the participants are presented with only one pattern in each trial, the strength of the
14 sound-shape correspondences can be better elucidated (e.g., Chen et al., 2019). Using such a single
15 pattern procedure, the mapping between “Kiki” and spiky patterns is typically more reliable than
16 that observed between “Bouba” and smooth patterns (as in the present study, Chen et al., 2016;
17 Nielsen & Rendall, 2011; Woods et al., 2013). Our conjecture is that when participants do not
18 perceive the pattern to be a good, simple pattern, they will tend to match it to “Kiki”. That is, the
19 variabilities in the perceptual grouping process within a particular participant group might
20 determine the consensuality of “Bouba” responses, but that of the “Kiki” responses was immune. It

1 is therefore interesting to investigate whether different perceptual styles (e.g., holistic vs. analytic)
2 would explain any individual or cultural differences in the Bouba/Kiki effect in the future studies
3 (e.g., Chen et al., 2016).

4 Our results, showing that the Bouba/Kiki effect is underpinned by perceptual grouping process,
5 provides critical implications for the development and neural mechanisms underpinning this
6 phenomenon. Even though rudimentary forms of the Bouba/Kiki effect have been observed in early
7 infancy (Ozturk et al., 2013; Pejovic & Molnar, 2017; Walker et al., 2010), the group-level
8 consensus of this sound-shape mapping develops continuously until late childhood (Chow &
9 Ciaramitaro, 2019; Tzeng et al., 2017). This protracted developmental trajectory should partly
10 depend on the visual development of shape processing (e.g., Cribb et al., 2016). Furthermore, the
11 core center of shape perception is currently suggested to be located at V4 where neurons have larger
12 receptive fields that can pool local segments in various orientations and spatial frequencies into a
13 global contour (Gallant et al., 1996; Wilkinson et al., 2000). Hence, we would like to suggest that
14 those brain areas that are sensitive to the congruency of the Bouba/Kiki effect should start from V4
15 upwards in the ventral stream of visual pathway. A recent human neuroimaging study of the
16 Bouba/Kiki effect demonstrated that the activities in the ventral occipitotemporal regions associated
17 with visual processing were mildly modulated by the congruency of the sounds (Peiffer-Smadja &
18 Cohen, 2019). Our results should provide a more focal region of interest for future imaging studies
19 of the Bouba/Kiki effect.

1 The Bouba/Kiki effect is suggested to be a basis for early language learning (Imai et al., 2015;
2 Monaghan et al., 2012; see Fort et al., 2018, for a review) and useful knowledge in product design
3 and marketing (Abel & Glinert, 2008; Chen et al., 2018; Ngo et al., 2011; Ngo et al. 2013; see
4 Spence, 2012, for a review). Another application is the sensory substitution device for the blind (see
5 Bach-y-Rita & Kercel, 2003; Deroy & Auvray, 2012, for reviews). Earlier vision-to-audition
6 sensory substitution devices typically decompose visual images into visual features and convert
7 them into auditory features using the already-known crossmodal correspondences (such as visual
8 elevation and auditory pitch, Capelle et al., 1998; Hamilton-Fletcher et al., 2016; Hanneton et al.,
9 2010; Meijer, 1992). The mid-level crossmodal correspondences that we have demonstrated here
10 might provide an alternative vision-to-audition coding scheme in that the complex visual shapes can
11 be directly converted into speech sounds.

12 **4.1 Consistency between the laboratory and online study**

13 We used an online study in Experiment 1 and a laboratory study in Experiment 2. The
14 advantage of online studies is that a large number of participants can be recruited in a short period
15 of time, and therefore the between-participant variables can be minimized; however, it is not
16 possible to control the parameters of stimulus presentation in online studies as in laboratory studies.
17 Nowadays, as the number of online participant pools and studies increase, researchers have also
18 demonstrated that the data obtained from online studies can be similarly reliable as laboratory
19 studies (Germine et al., 2012; Pauszek et al., 2017; Woods et al., 2015). Therefore, convergent
20 evidence from online and laboratory studies can hopefully strengthen the robustness and

1 generalization of the conclusions. In the Bouba/Kiki effect specifically, Chen et al. (2016)
2 demonstrated that this effect is insensitive to the size of visual patterns, and similar results can be
3 observed in online and laboratory studies. In the current study, the modulation of the curvature
4 factor (convexities or concavities) on the Bouba/Kiki effect were reliably observed when using
5 either method. We therefore suggest that online studies provide useful and reliable tools with which
6 to examine the multi-faceted of the Bouba/Kiki effect and crossmodal correspondences.

7

8 **5. Conclusion**

9 Here we demonstrated that the Bouba/Kiki effect, specifically the patterns matched to “Bouba”
10 should be underpinned by the shape perception rather than feature level of visual information
11 processing. In future research it will be intriguing to examine whether any corresponding level in
12 the auditory information processing underpinning the Bouba/Kiki effect as well. Our results
13 provides critical implications to the development and neural mechanisms of the Bouba/Kiki effect,
14 as well as the possible applications in the fields of language learning, marketing, and sensory
15 rehabilitations.

16

Acknowledgments

YCC and CS were supported by the Arts and Humanities Research Council (AHRC), Rethinking the Senses grant (AH/L007053/1). YCC is supported by Ministry of Science and Technology in Taiwan (MOST 107-2410-H-715-001-MY2). PCH is supported by Ministry of Science and Technology in Taiwan (MOST 104-2628-H-006-001-MY3). We would like to thank Dr. Andy Woods for his help in building the online study in Experiment 1.

References

- Abel, G. A., & Glinert, L. H. (2008). Chemotherapy as language: Sound symbolism in cancer medication names. *Social Science & Medicine*, 66, 1863-1869.
- 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.
- Bach-y-Rita, P., & Kercel, S. W. (2003). Sensory substitution and the human-machine interface. *Trends in Cognitive Sciences*, 7, 541-546.
- Baldwin, A. S., Schmidtmann, G., Kingdom, F. A. A., & Hess, R. F. (2016), Rejecting probability summation for radial frequency patterns, not so quick! *Vision Research*, 122, 124-134.
- Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., ... Grothendieck, G. (2015). *Linear Mixed-Effects Models using 'Eigen' and S4*. Retrieved from <https://cran.r-project.org/web/packages/lme4/lme4.pdf> (on 12/10/2015).
- Bell, J., Hancock, S., Kingdom, F. A. A., & Peirce, J. W. (2010). Global shape processing: Which parts form the whole? *Journal of Vision*, 10(6):16, 1-13. doi:10.1167/10.6.16.

- 1 Berlin, B. (1994). Evidence for pervasive synesthetic sound symbolism in ethnozoological
2 nomenclature. In L. Hinton, J. Nichols, & J. Ohala (Eds.), *Sound symbolism* (pp. 76-93). New
3 York, NY: Cambridge University Press.
- 4 Bottini, R., Barilari, M., & Collignon, O. (2019). Sound symbolism in sighted and blind. The role of
5 vision and orthography in sound-shape correspondences. *Cognition*, 185, 62-70.
- 6 Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10, 433-436.
- 7 Bremner, A., Caparos, S., Davidoff, J., de Fockert, J., Linnell, K., & Spence, C. (2013). Bouba and
8 Kiki in Namibia? A remote culture make similar shape-sound matches, but different shape-taste
9 matches to Westerners. *Cognition*, 126, 165-172.
- 10 Capelle, C., Trullemans, C., Arno, P., & Veraart, C. (1998). A real-time experimental prototype for
11 enhancement of vision rehabilitation using auditory substitution. *IEEE Transactions on*
12 *Biomedical Engineering*, 45, 1279-1293.
- 13 Chater, N., & Vitányi, P. (2003). Simplicity: A unifying principle in cognitive science? *Trends in*
14 *Cognitive Sciences*, 7, 19-22.
- 15 Chen, Y.-C., Huang, P.-C., Woods, A., & Spence, C. (2016). When “Bouba” equals “Kiki”:
16 Cultural commonalities and cultural differences in sound-shape correspondences. *Scientific*
17 *Reports*, 6:26681.
- 18 Chen, Y.-C., Huang, P.-C., Woods, A., & Spence, C. (2019). I know that “Kiki” is angular: The
19 metacognition underlying sound–shape correspondences. *Psychonomic Bulletin & Review*, 26,
20 261-268.
- 21 Chen, Y.-C., & Spence, C. (2017). Assessing the role of the ‘unity assumption’ on multisensory
22 integration: A review. *Frontiers in Psychology*, 8:445.
23 <https://doi.org/10.3389/fpsyg.2017.00445>
- 24 Chen, Y.-C., Woods, A., & Spence, C. (2018). Sensation transference from plateware to food: The
25 sounds and tastes of plates. *International Journal of Food Design*, 3, 41-62.
- 26 Chow, H. M., & Ciaramitaro, V. (2019). What makes a shape “baba”? The shape features
27 prioritized in sound–shape correspondence change with development. *Journal of Experimental*
28 *Child Psychology*, 179, 73-89.

- 1 Cribb, S. J., Badcock, J. C., Maybery, M. T., & Badcock, D. R. (2016). Dissociation of local and
2 global contributions to detection of shape with age. *Journal of Experimental Psychology:*
3 *Human Perception and Performance*, 42, 1761-1769.
- 4 Davis, R. (1961). The fitness of names to drawings. A cross-cultural study in Tanganyika. *British*
5 *Journal of Psychology*, 52, 259-268.
- 6 Deroy, O., & Auvray, M. (2012). Reading the world through the skin and ears: A new perspective
7 on sensory substitution. *Frontiers in Psychology*, 3:457.
8 <https://doi.org/10.3389/fpsyg.2012.00457>
- 9 Fantoni, C., Bertamini, M., & Gerbino, W. (2005). Contour curvature polarity and surface
10 interpolation. *Vision Research*, 45, 1047-1062.
- 11 Fantoni, C., & Gerbino, W. (2003). Contour interpolation by vector-field combination. *Journal of*
12 *Vision*, 3, 281-303. doi:<https://doi.org/10.1167/3.4.4>
- 13 Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. (2007). G*Power 3: A flexible statistical power
14 analysis program for the social, behavioral, and biomedical sciences. *Behavior Research*
15 *Methods*, 39, 175-191.
- 16 Feldman, J. (1999). The role of objects in perceptual grouping. *Acta Psychologica*, 102, 137-163.
- 17 Feldman, J. (2009). Bayes and the simplicity principle in perception. *Psychological Review*, 116,
18 875-887.
- 19 Feldman, J., & Singh, M. (2005). Information along contours and object boundaries. *Psychological*
20 *Review*, 112, 243-252.
- 21 Fort, M., Lammertink, I., Peperkamp, S., Guevara-Rukoz, A., Fikkert, P., & Tsuji, S. (2018).
22 Symbouki: A meta-analysis on the emergence of sound symbolism in early language
23 acquisition. *Developmental Science*, 21:e12659.
- 24 Froyen, V., Feldman, J., & Singh, M. (2015). Bayesian hierarchical grouping: Perceptual grouping
25 as mixture estimation. *Psychological Review*, 122, 575-597.
- 26 Fryer, L., Freeman, J., & Pring, L. (2014). Touching words is not enough: How visual experience
27 influences haptic–auditory associations in the “Bouba–Kiki” effect. *Cognition*, 132, 164-173.

- 1 Gallant, J. L., Connor, C. E., Rakshit, S., Lewis, J. W., & Van Essen, D. C. (1996). Neural
2 responses to polar, hyperbolic, and Cartesian gratings in area V4 of the macaque monkey.
3 *Journal of Neurophysiology*, 76, 2718-2739.
- 4 Germine, L., Nakayama, K., Duchaine, B. C., Chabris, C. F., Chatterjee, G., & Wilmer, J. B.
5 (2012). Is the Web as good as the lab? Comparable performance from Web and lab in
6 cognitive/perceptual experiments. *Psychonomic Bulletin & Review*, 19, 847-857.
- 7 Hamilton-Fletcher, G., Pisanski, K., Reby, D., Stefańczyk, M., Ward, J., & Sorokowska, A. (2018).
8 The role of visual experience in the emergence of cross-modal correspondences. *Cognition*,
9 175, 114-121.
- 10 Hamilton-Fletcher, G., Wright, T. D., & Ward, J. (2016). Cross-modal correspondences enhance
11 performance on a colour-to-sound sensory substitution device. *Multisensory Research*, 29, 337-
12 363.
- 13 Hanneton, S., Auvray, M., & Durette, B. (2010). The Vibe: A versatile vision-to-audition sensory
14 substitution device. *Applied Bionics and Biomechanics*, 7, 269-276.
- 15 Hess, R. F., Wang, Y. Z., & Dakin, S. C. (1999). Are judgements of circularity local or global?
16 *Vision Research*, 39, 4354-4360.
- 17 Holland, M. K., & Wertheimer, M. (1964). Some physiognomic aspects of naming, or, maluma and
18 takete revisited. *Perceptual and Motor Skills*, 19, 111-117.
- 19 Hubel, D. H. & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional
20 architecture in the cat's visual cortex. *Journal of Physiology*, 160, 106-154.
- 21 Imai, M., & Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language
22 acquisition and language evolution. *Philosophical Transactions of the Royal Society B:*
23 *Biological Sciences*, 369(1651):20130298.
- 24 Imai, M., Miyazaki, M., Yeung, H. H., Hidaka, S., Kantartzis, K., Okada, H., & Kita, S. (2015).
25 Sound symbolism facilitates word learning in 14-month-olds. *PLoS One*, 10:e0116494.
26 <https://doi.org/10.1371/journal.pone.0116494>
- 27 Knoeferle, K., Li, J., Maggioni, E., & Spence, C. (2017). What drives sound symbolism? Different
28 acoustic cues underlie sound-size and sound-shape mappings. *Scientific Reports*, 7(1):5562.

- 1 Köhler, W. (1910). Akustische Untersuchungen [Acoustic investigations]. *Zeitschrift für*
2 *Psychologie*, 2, 59-140.
- 3 Köhler, W. (1929). *Gestalt psychology*. New York, NY: Liveright.
- 4 Kubovy, M., & Van Valkenburg, D. (2001). Auditory and visual objects. *Cognition*, 80, 97-126.
- 5 Lewkowicz, D. J., & Minar, N. J. (2014). Infants are not sensitive to synesthetic cross-modality
6 correspondences: A comment on Walker et al. (2010). *Psychological Science*, 25, 832-834.
- 7 Loffler, G., Wilson, H. R., & Wilkinson, F. (2003). Local and global contributions to shape
8 discrimination. *Vision Research*, 43, 519-530.
- 9 Marks, L. E. (1987). On cross-modal similarity: Auditory-visual interactions in speeded
10 discrimination. *Journal of Experimental Psychology: Human Perception and Performance*, 13,
11 384-394.
- 12 Marks, L. E. (1996). On perceptual metaphors. *Metaphor and Symbolic Activity*, 11, 39-66.
- 13 Marks, L. E. (2013). Weak synesthesia in perception and language. In J. Simner & E. Hubbard
14 (Eds.), *Oxford handbook of synesthesia* (pp. 761-789). Oxford, UK: Oxford University Press.
- 15 Maurer, D., Pathman, T., & Mondloch, C. J. (2006). The shape of boubas: Sound-shape
16 correspondences in toddlers and adults. *Developmental Science*, 9, 316-322.
- 17 Meijer, P. B. (1992). An experimental system for auditory image representations. *IEEE*
18 *Transactions on Biomedical Engineering*, 39, 112-121.
- 19 Monaghan, P., Mattock, K., & Walker, P. (2012). The role of sound symbolism in language
20 learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 1152-
21 1164.
- 22 Ngo, M. K., Misra, R., & Spence, C. (2011). Assessing the shapes and speech sounds that people
23 associate with chocolate samples varying in cocoa content. *Food Quality and Preference*, 22,
24 567-572.
- 25 Ngo, M. K., Velasco, C., Salgado, A., Boehm, E., O'Neill, D., & Spence, C. (2013). Assessing
26 crossmodal correspondences in exotic fruit juices: The case of shape and sound
27 symbolism. *Food Quality and Preference*, 28, 361-369.

- 1 Nielsen, A., & Rendall, D. (2011). The sound of round: Evaluating the sound-symbolic role of
2 consonants in the classic Takete-Maluma phenomenon. *Canadian Journal of Experimental*
3 *Psychology*, 65, 115-124.
- 4 Nielsen, A. K., & Rendall, D. (2013). Parsing the role of consonants versus vowels in the classic
5 Takete-Maluma phenomenon. *Canadian Journal of Experimental Psychology*, 67, 153-163.
- 6 Nuckolls, J. B. (1999). The case for sound symbolism. *Annual Review of Anthropology*, 28, 225-
7 252.
- 8 Ozturk, O., Krehm, M., & Vouloumanos, A. (2013). Sound symbolism in infancy: Evidence for
9 sound-shape cross-modal correspondences in 4-month-olds. *Journal of Experimental Child*
10 *Psychology*, 114, 173-186.
- 11 Parise, C. V., & Spence, C. (2009). ‘When birds of a feather flock together’: Synesthetic
12 correspondences modulate audiovisual integration in non-synesthetes. *PLoS One*, 4(5):e5664.
13 <https://doi.org/10.1371/journal.pone.0005664>
- 14 Parise, C. V., Spence, C., & Ernst, M. (2012). When correlation implies causation in multisensory
15 integration. *Current Biology*, 22, 46-49.
- 16 Pauszek, J. R., Sztybel, P., & Gibson, B. S. (2017). Evaluating Amazon’s Mechanical Turk for
17 psychological research on the symbolic control of attention. *Behavior Research Methods*, 49,
18 1969-1983.
- 19 Peiffer-Smadja, N., & Cohen, L. (2019). The cerebral bases of the bouba-kiki effect. *NeuroImage*,
20 186, 679-689.
- 21 Pejovic, J., & Molnar, M. (2017). The development of spontaneous sound-shape matching in
22 monolingual and bilingual infants during the first year. *Developmental Psychology*, 53, 581-
23 586.
- 24 Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia – A window into perception, thought
25 and language. *Journal of Consciousness Studies*, 8, 3-34.
- 26 Ramachandran, V. S., Marcus, Z., & Chunharas, C. (2020). Bouba-Kiki: Cross-domain resonance
27 and the origins of synesthesia, metaphor, and words in the human mind. In K. Sathian & V. S.

Ramachandran (Eds.), *Multisensory perception: From laboratory to clinic* (pp. 3-40). San Diego, CA: Elsevier.

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, 2, 1019-1025.

Rogers, S. K., & Ross, A. S. (1975). A cross-cultural test of the Maluma-Takete phenomenon. *Perception*, 4, 105-106.

Rohe, T., Ehrls, A. C., & Noppeney, U. (2019). The neural dynamics of hierarchical Bayesian causal inference in multisensory perception. *Nature Communications*, 10:1907.
<https://doi.org/10.1038/s41467-019-09664-2>

Rohe, T., & Noppeney, U. (2015). Cortical hierarchies perform Bayesian causal inference in multisensory perception. *PLoS Biology*, 13:e1002073. doi:10.1371/journal.pbio.1002073

Schmidtman, G., Jennings, B. J., & Kingdom, F. A. (2015). Shape recognition: Convexities, concavities and things in between. *Scientific Reports*, 5:17142.

Schmidtman, G., & Fruend, I. (2019). Radial frequency patterns describe a small and perceptually distinct subset of all possible planar shapes. *Vision Research*, 154, 122-130,
<https://doi.org/10.1016/j.visres.2018.10.007>.

Schmidtman, G., & Kingdom, F. A. A. (2017). Nothing more than a pair of curvatures: A common mechanism for the detection of both radial and non-radial frequency patterns. *Vision Research*, 134, 18-25,

Serre, T., Oliva, A., Poggio, T. (2007). A feedforward architecture accounts for rapid categorization. *Proceedings of the National Academy of Sciences of the USA*, 104, 6424-6429.

Shams, L., & Beierholm, U. R. (2010). Causal inference in perception. *Trends in Cognitive Sciences*, 14, 425-432.

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*, 8:2139. doi: 10.3389/fpsyg.2017.02139

Sidhu, D. M., & Pexman, P. M. (2018). Five mechanisms of sound symbolic association. *Psychonomic Bulletin & Review*, 25, 1619-1643.

- 1 Sievers, B., Lee, C., Haslett, W., & Wheatley, T. (2019). A multi-sensory code for emotional
2 arousal. *Proceedings of the Royal Society B*, 286, 20190513.
3 <https://doi.org/10.1098/rspb.2019.0513>
- 4 Silva, D. M. R., & Bellini-Leite, S. C. (2019). Cross-modal correspondences in sine wave: Speech
5 versus non-speech modes. *Attention, Perception, & Psychophysics*.
6 <https://doi.org/10.3758/s13414-019-01835-z>
- 7 Sourav, S., Kekunnaya, R., Shareef, I., Banerjee, S., Bottari, D., & Röder, B. (2019). A protracted
8 sensitive period regulates the development of cross-modal sound–shape associations in
9 humans. *Psychological Science*, 30, 1473-1482.
- 10 Spector, F., & Maurer, D. (2013). Early sound symbolism for vowel sounds. *i-Perception*, 4, 239-
11 241. <https://doi.org/10.1068/i0535>
- 12 Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, &*
13 *Psychophysics*, 73, 971-995.
- 14 Spence, C. (2012). Managing sensory expectations concerning products and brands: Capitalizing on
15 the potential of sound and shape symbolism. *Journal of Consumer Psychology*, 22, 37-54.
- 16 Spence, C., & Chen, Y.-C. (2012). Intramodal and crossmodal perceptual grouping. In B. Stein
17 (Ed), *The new handbook of multisensory processing* (pp. 265-282). Cambridge, MA: MIT
18 Press.
- 19 Spence, C., & Frings, C. (2019). Multisensory feature integration in (and out) of the focus of spatial
20 attention. *Attention, Perception, & Psychophysics*, 82(1), 363-376.
21 <https://doi.org/10.3758/s13414-019-01813-5>
- 22 Spence, C., & Sathian, K. (2020). Audiovisual crossmodal correspondences: Behavioural
23 consequences & neural underpinnings. In K. Sathian & V. S. Ramachandran (Eds.),
24 *Multisensory perception: From laboratory to clinic* (pp. 239-258). San Diego, CA: Elsevier.
- 25 Spence, C., Wan, X., Woods, A., Velasco, C., Deng, J., Youssef, J., & Deroy, O. (2015). On tasty
26 colours and colourful tastes? Assessing, explaining, and utilizing crossmodal correspondences
27 between colours and basic tastes. *Flavour*, 4:23. <https://doi.org/10.1186/s13411-015-0033-1>
- 28 Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses*. Cambridge, MA: MIT Press.

- 1 Todorovic, D. (2008). Gestalt principles. *Scholarpedia*, 3(12):5345.
- 2 Tzeng, C. Y., Nygaard, L. C., & Namy, L. L. (2017). Developmental change in children's
3 sensitivity to sound symbolism. *Journal of Experimental Child Psychology*, 160, 107-118.
- 4 Van Essen, D.C., & Gallant, J.L. (1994). Neural mechanisms of form and motion processing in the
5 primate visual system. *Neuron*, 13, 1-10.
- 6 Wagemans, J., Feldman, J., Gepshtein, S., Kimchi, R., Pomerantz, J. R., Van der Helm, P. A., &
7 Van Leeuwen, C. (2012). A century of Gestalt psychology in visual perception: II. Conceptual
8 and theoretical foundations. *Psychological Bulletin*, 138, 1218-1252.
- 9 Walker, P. (2012). Cross-sensory correspondences and cross talk between dimensions of
10 connotative meaning: Visual angularity is hard, high-pitched, and bright. *Attention, Perception,*
11 *& Psychophysics*, 74, 1792-1809.
- 12 Walker, P., Bremner, J. G., Mason, U., Spring, J., Mattock, K., Slater, A., & Johnson, S. P. (2010).
13 Preverbal infants' sensitivity to synaesthetic cross-modality correspondences. *Psychological*
14 *Science*, 21, 21-25.
- 15 Walker, P., Bremner, J. G., Mason, U., Spring, J., Mattock, K., Slater, A., & Johnson, S. P. (2014).
16 Preverbal infants are sensitive to cross-sensory correspondences: Much ado about the null
17 results of Lewkowicz and Minar (2014). *Psychological Science*, 25, 835-836.
- 18 Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy.
19 *Psychological Bulletin*, 3, 638-667.
- 20 Wertheimer, M. (1938). Laws of organization in perceptual forms. In W. D. Ellis (Ed.), *A source*
21 *book of Gestalt psychology* (pp. 71-88). London, UK: Kegan Paul, Trench, Trubner &
22 Company.
- 23 Westbury, C., Hollis, G., Sidhu, D. M., & Pexman, P. M. (2018). Weighing up the evidence for
24 sound symbolism: Distributional properties predict cue strength. *Journal of Memory and*
25 *Language*, 99, 122-150.
- 26 Wilkinson, F., James, T. W., Wilson, H. R., Gati, J. S., Menon, R. S., & Goodale, M. A. (2000). An
27 fMRI study of the selective activation of human extrastriate form vision areas by radial and
28 concentric gratings. *Current Biology*, 10, 1455-1458.

- 1 Wilkinson, F., Wilson, H. R., & Habak, C. (1998). Detection and recognition of radial frequency
2 patterns. *Vision Research*, 38, 3555-3568.
- 3 Wilson, H. R., Loffler, G., Wilkinson, F. (2002). Synthetic faces, face cubes, and the geometry of
4 face space. *Vision Research*, 42, 2909-2923. [https://doi.org/10.1016/S0042-6989\(02\)00362-0](https://doi.org/10.1016/S0042-6989(02)00362-0)
- 5 Wilson, H. R., & Wilkinson, F. (2002). Symmetry perception: A novel approach for biological
6 shapes. *Vision Research*, 42, 589-597. [https://doi.org/10.1016/S0042-6989\(01\)00299-1](https://doi.org/10.1016/S0042-6989(01)00299-1)
- 7 Woods, A. T., & Spence, C. (2016). Using single colors and color pairs to communicate basic
8 tastes. *i-Perception*, 7:4. <https://doi.org/10.1177/2041669516658817>
- 9 Woods, A. T., Spence, C., Butcher, N., & Deroy, O. (2013). Fast lemons and sour boulders: Testing
10 crossmodal correspondences using an internet-based testing methodology. *i-Perception*, 4, 365-
11 379. <https://doi.org/10.1068/i0586>
- 12 Woods, A. T., Velasco, C., Levitan, C. A., Wan, X., & Spence, C. (2015). Conducting perception
13 research over the internet: A tutorial review. *Peer J*, 3:e1058.
14 <https://doi.org/10.7717/peerj.1058>

Appendix

To create unfamiliar but object-like contours, we used compound radial frequency (RF) patterns in the current study (Schmidtman et al., 2015). We summed three different RF components, which were defined as follows:

$$r_{compound}(\theta) = r_{mean}(1 + A_1 \sin(\omega_1 \theta + \varphi_1) + A_2 \sin(\omega_2 \theta + \varphi_2) + A_3 \sin(\omega_3 \theta + \varphi_3)) \quad \text{Eq.(A.1)}$$

In which r_{mean} is the radius of the base circle, A is the modulation amplitude, ω is the radial frequency, and φ is phase of each RF component, respectively. We chose radial frequencies of 3, 5, and 8. Six different amplitudes combinations (A_1, A_2, A_3) were tested in which their sum equals 0.5: (0.1, 0.2, 0.2), (0.2, 0.1, 0.2), (0.2, 0.2, 0.1), (0.3, 0.1, 0.1), (0.1, 0.3, 0.1), and (0.1, 0.1, 0.3). Four phase combinations ($\varphi_1, \varphi_2, \varphi_3$) were used: (-0.5, 0.5, 0), (0.5, -0.5, 0), (-0.25, 0.25, 0) and (0.25, -0.25, 0). The manipulation of the phase produced the compound RF patterns having different rotation angles.

The segment length of the segmental RF patterns was defined as the percentage of the angles that the compound RF pattern covered. When producing eight segments, the center of convexity and concavity was based on that of RF8. While when producing five segments, the center of convexity and concavity was based on that of RF5. The center of each segment corresponds to the peak of each convexity or the valley of each concavity. The reverse of the segments was created using the software Photoshop. Each segment was rotated 180 degrees along the radial axis.

Schmidtman et. al (2015) have demonstrated that the segments of convexities play a more critical role than concavities in recognizing the closed curvilinear shapes. In order to investigate whether the stimuli used in Experiment 1 also exhibited such a superiority for convexities over concavities when grouping these segments to construct a pattern, we adapted the match-to-sample task (Schmidtman et al., 2015).

Twenty-Five participants who were naïve to the purpose of the study and did not participate Experiments 1 and 2 took part in this study. The apparatus was the same as for Experiment 2. Two factors were manipulated: curvature (convexities and concavities) and segment length (25%, 50%, and 75%) (same as in Experiment 1). Thirty-six segmental RF patterns (6 amplitude combinations x

2 curvature x 3 segment length) were created. Each segmental RF pattern was tested 30 times, giving rise to a total of 1080 trials. The experiment took about 35 minutes to complete.

Each trial consisted of a 400 ms presentation of a segmental RF pattern as reference shape, followed by a 300 ms blank interval, and then the simultaneous side-by-side presentation of two compound RF pattern as target and distractor shapes. The target was the complete pattern of the segmental RF pattern, and the distractors were randomly chosen from other compound RF patterns that had been used in this experiment. The location of the target (on the left or right side of the display) was randomly determined on a trial-by-trial basis. The target and distractors were presented until the participants responded. In order to avoid the possibilities that participants used local cues (e.g. distance between local features) to complete the matching task, the size of the target and distractors were scaled by a random amount between $-1/2$ to $+1/4$ in size relative to the reference shape (Schmidtman et al. 2015). The participants had to match a segmental RF pattern to one of two compound RF patterns (one target and the other distractor).

The results are shown in Figure A1. We averaged the accuracy for all segmental RF patterns in each participant. The mean accuracy across 25 participants was plotted against the segment length. The squares represent the results for concavities, and the circles represent the results for convexities. The error bars indicate ± 1 standard error of the mean. The mean accuracy data was submitted to a two-way repeated-measure ANOVA. The main effect of curvature was significant ($F(1,24) = 4.81, p < .05$), where the accuracy was higher for convexities than concavities. The main effect of segment length was also significant ($F(2,48) = 231.00, p < .001$). However, their interaction was not significant ($F(2,48) = 1.44, p = .25$). These result therefore demonstrates that the segments of convexities were easier to group and reconstruct into original compound RF patterns as compared to concavities, thus replicating Schmidtman et al.'s (2015) results.

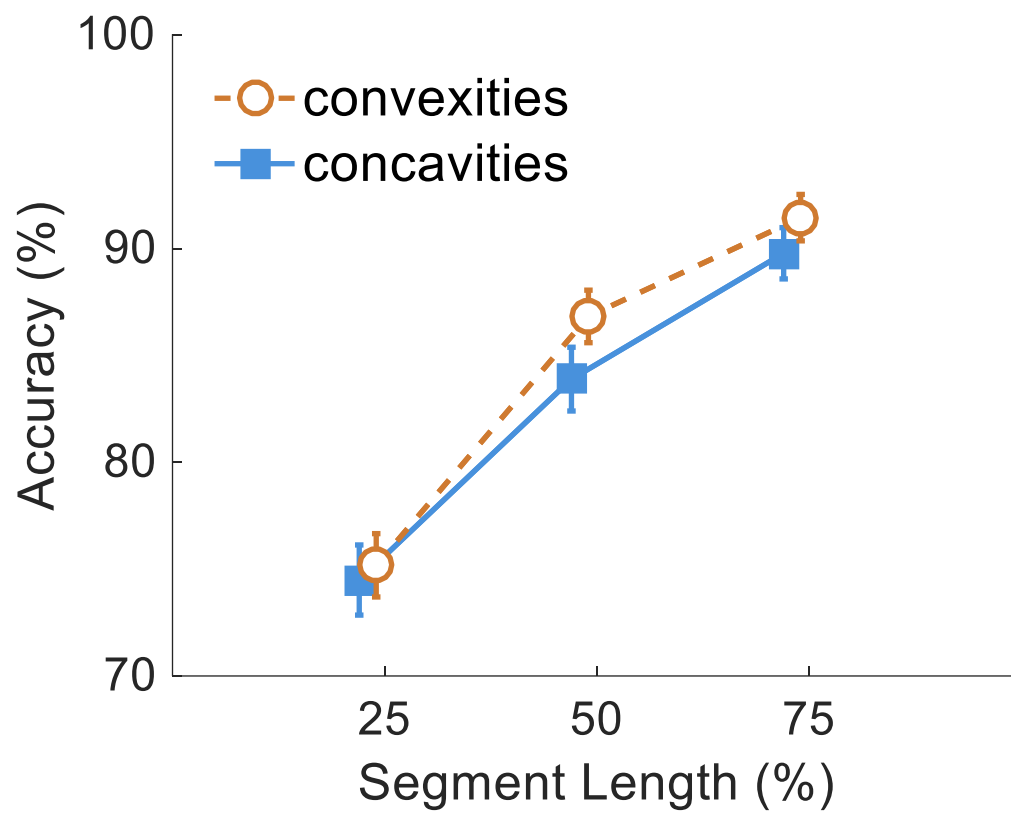


Figure A1. Mean accuracy of matching responses for segmental and compound RF patterns. The open circles and orange dashed line represent the results for convexities, and the solid squares and blue solid line represent the results for concavities. Error bars represent ± 1 standard error (SE) of the mean.