Journal of Experimental Child Psychology 179 (2019) 73–89

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

What makes a shape "baba"? The shape features prioritized in sound–shape correspondence change with development

Hiu Mei Chow, Vivian Ciaramitaro *

Department of Psychology, University of Massachusetts Boston, Boston, MA 02125, USA

ARTICLE INFO

Article history:
Received 12 March 2018
Revised 15 October 2018

Keywords:
Crossmodal sound–shape correspondence
Sound symbolism
Developmental change
Perceptual processing
Contextual comparison
Bouba-kiki effect

ABSTRACT

Crossmodal sound–shape correspondence, the association of abstract shapes and nonsense words (e.g., "bouba-kiki" effect), is seen across cultures and languages. Recent research suggests that the sensitivity for such sound–shape pairings might increase with development. Here we examined one possible mechanism underlying developmental changes in sensitivity to sound–shape correspondences—if and how shape features, such as contour spikiness and the number and size of protrusions, might be weighted differently by children and adults. In Experiment 1, we asked participants to choose which of two nonsense words matched a given visual shape while manipulating contour spikiness and number and size of protrusions independently. We found that adults associated /i/ sounds with shapes having spiky contours and 3 small protrusions. Of these shape features, contour spikiness showed the strongest association. Whereas 9- to 11-year-olds showed adult-like responses, 6- to 8-year-olds prioritized protrusion number, not contour spikiness. Importantly, in Experiment 2, where contour spikiness was highlighted by presenting round and spikey shapes side by side, 6- to 8-year-olds could make associations based on contour spikiness. Our findings suggest that 6- to 8-year-olds prioritize different features of a shape when making sound–shape correspondence compared with adults. Interestingly, these shape-processing biases can be altered by context such that children can resemble adults when the relevant shape features are highlighted. Our results suggest that biases in visual shape processing and the ability to extract contextual information might be additional factors explaining developmental

* Corresponding author.
E-mail address: vivian.ciaramitaro@umb.edu (V. Ciaramitaro).

https://doi.org/10.1016/j.jecp.2018.10.005

changes in sensitivity toward sound–shape correspondences. These changing developmental biases highlight the contribution of perceptual processing styles in crossmodal correspondence.

Introduction

Efficiently operating in the world involves integrating inputs from our different senses. What information do we put together? Often information from different senses originates from a common source, conveying redundant information that we tend to integrate (e.g., a face and voice both convey happiness). However, different sensory experiences may be associated even if not originating from a common source, reflecting a consistent crossmodal correspondence across individuals (reviewed by Spence, 2011). For example, in the "bouba-kiki" effect, first reported by Köhler (1929, 1947) and since replicated numerous times, a spikey shape tends to be associated with speech sounds like /kiki/ and /takete/, whereas a rounded shape tends to be associated with speech sounds like /bouba/ and /maluma/. Studies have addressed the underlying neural substrates for sound–shape correspondence (e.g., Ramachandran & Hubbard, 2003; Spence & Parise, 2012), its role on language acquisition (e.g. Imai & Kita, 2014), and the evolution of language on sound–shape correspondences (Maurer & Mondloch, 2004) as well as practical implications for enhanced learning (e.g., Imai et al., 2015), advertising (e.g., Belli & Sagrillo, 2001), and sensory substitutions (e.g., Hamilton-Fletcher, Wright, & Ward, 2015; Stiles & Shimojo, 2015). Here we explored how sound–shape correspondences develop across childhood and factors contributing to this development.

Sound–shape correspondences in adults

Sound–shape correspondence is robust across many languages. Whereas the bouba-kiki effect was first established in Spanish speakers (Köhler 1929, 1947), it has been found in other spoken languages as well (e.g., Nielsen & Rendall, 2011; Ramachandran & Hubbard, 2003; Westbury, 2005) and in individuals without written language (e.g., Bremner et al., 2013). Sound–shape correspondence is also robust across testing paradigms such as label matching (e.g., Köhler, 1929), speeded classification (e.g., Lupyan & Casasanto, 2015; Westbury, 2005), word generation (e.g., Nielsen & Rendall, 2013), and adaptation paradigms (Sweeny, Guzman-Martinez, Ortega, Grabowecky, & Suzuki, 2012). Despite the robustness of the bouba-kiki effect across languages and methods, several factors are known to influence sound–shape correspondence and its development.

For instance, an adult's perceptual processing style can modulate correspondences. Shape features such as contour spikiness and the number and size of protrusions are important visual dimensions influencing the bouba-kiki effect (Chen, Huang, Woods, & Spence, 2016). Using radial frequency patterns and independently manipulating parameters such as frequency cycle (number of protrusions), amplitude (size of protrusions), and number of triangular harmonics added to the sinusoidal wave forming the pattern, with more harmonics creating a rougher-looking pattern (contour spikiness), Chen et al. (2016) found that participants were more likely to choose a /kiki/ sound when all three parameters of the radial frequency pattern increased. In addition, the weighting of a particular visual feature varied based on whether participants were from a North American or East Asian culture. North Americans (traditionally considered a more individualistic culture) placed more weight on amplitude, whereas East Asians (traditionally considered a more holistic culture) placed more weight on contour spikiness. The authors speculated that cultural differences were rooted in perceptual processing style differences, with different cultures differentially prioritizing global features (focusing on patterns emerging from combining individual lobes) versus local features (focusing on individual lobes).

Developmental changes of sound–shape correspondences

In terms of development, sound–shape correspondence has been found early in development. Infants as young as 4 months exhibit a bouba-kiki sound–shape correspondence (Ozturk, Krehm, & Vouloumanos, 2013). Older children also exhibit sound–shape correspondence (11- to 14-month-olds: Asano et al., 2015; Pejovic & Molnar, 2017;2.5- to 3-year-olds: Maurer, Pathman, & Mondloch, 2006; Spector & Maurer, 2013; 3- to 7-year-olds: Tzeng, Nygaard, & Namy, 2017; 8- and 9-year-olds or above: Davis, 1961; Irwin & Newland, 1940).

Although sound–shape correspondence has been found in many age groups, it is not necessarily robust early in development. First, a recent meta-analysis across participants ranging in age from 4 months to 3 years showed that the effect size of sound–shape correspondence was rather small at 0.12 (95% confidence interval [−0.02, 0.25]), as estimated from a random effects model (Fort et al., 2018; see Styles & Gawne, 2017, for a meta-analysis in adults). Second, the *strength* of sound–shape correspondences has been shown to vary across development; a smaller proportion of 2.5-year-olds exhibited the association compared with college-age adults (Maurer et al., 2006). Furthermore, 3-year-olds were not able to generalize sound–shape correspondence to foreign words, unlike 5- to 7-year-olds who were (Tzeng et al., 2017), suggesting weaker sensitivity to sound–shape correspondence in the younger age group. Third, 4-month-olds who showed the association exhibited it only for a combination of consonant and vowel sounds (/kiki/ vs. /bubu/) but not for only consonants (/kiki/ vs. /bibi/) or only vowels (/kiki/ vs. /kuku/), unlike adults who showed the association under all three conditions (Ozturk et al., 2013). Fourth, studies of other types of crossmodal correspondences have reported mixed results in terms of the early emergence of crossmodal correspondences (Lewkowicz & Minar, 2014; Walker et al., 2010, 2014). Altogether, evidence to date suggests that there could be both qualitative and quantitative changes in sound–shape correspondences across development, but the factors causing these changes are poorly understood.

Developmental changes of perceptual processing style

Developmental changes in the strength of sound–shape correspondence could be driven by changes in perceptual processing styles. More specifically, older children and adults may adopt different perceptual processing styles than younger children, which could yield qualitative differences in sound–shape correspondence across age groups. Navon-like hierarchical patterns in which a global pattern is formed by local elements, or Kanizsa triangle patterns in which a global illusory contour is formed by local Pacman-like inducers, are often used to probe global/local processing styles because global and local features can be manipulated independently. Despite inconsistencies across studies regarding the direction of change, many studies report changes in global/local processing styles across development within the same culture. Whereas some studies find a developmental switch from prioritizing local processing to prioritizing global processing (Kimchi, Hadad, Behrmann, & Palmer, 2005; Nayar, Franchak, Adolph, & Kiorpes, 2015; Oishi et al., 2014; Poirel et al., 2011; Poirel, Mellet, Houdé, & Pineau, 2008; Scherf, Behrmann, Kimchi, & Luna, 2009), other studies find a developmental switch from global to local processing styles (Kimchi et al., 2005; Kovshoff, Iarocci, Shore, & Burack, 2015). Still other studies find that children can use both levels of processing flexibly (Puspitawati, Jebrane, & Vinter, 2014). If perceptual processing style modulates crossmodal correspondence in adults (Chen et al., 2016), and younger children may adopt different processing styles than older children and adults, then the crossmodal correspondences found in younger children may differ from those found in older participants due to processing style differences. To address this possibility, we need a finer understanding of what visual shape features matter most to the bouba-kiki effect at different ages.

The current study

We examined qualitative and quantitative changes in sound–shape correspondence in development as a function of visual shape features. Across two experiments, we quantified the direction and magnitude of sound–shape associations in 6- to 35-year-olds using the same stimuli and similar paradigms. In Experiment 1 (two-interval forced choice, similar to Chen et al., 2016; Nielsen & Rendall, 2011), participants choose which of two nonsense words matched a shape. In Experiment 2 (two-alternative forced choice, similar to Maurer et al., 2006), participants chose which of two shapes matched a sound.

We used a novel set of shape stimuli and manipulated shape features including contour spikiness and number and size of protrusions, similar to the features manipulated in the radial frequency patterns used by Chen et al. (2016). We expected adults to show a significant association between nonsense words and visual shapes depending on shape features, extending previous findings (Chen et al., 2016). We conceptualized protrusion number as the most prominent global feature because determining the total number of protrusions requires inspecting the whole shape. We conceptualized contour spikiness and protrusion size as local features because determining either feature is possible by focusing only on individual protrusions. If crossmodal correspondence is influenced by perceptual processing style, young children might prioritize different features than adults based on whether features are global or local. However, given the mixed findings for developmental switches in perceptual processing styles, we had no a priori assumption regarding whether young children would prioritize local shape features (contour spikiness or protrusion size) or global shape features (protrusion number). Either way, changes in the features prioritized by children would illuminate qualitative and quantitative insights into the shape features important across development.

Experiment 1: Two-interval forced choice (sound bias without highlighting relevant feature)

Method

Participants

A total of 21 adults (18–35 years old) and 49 children (6–17 years old) participated in Experiment 1 (see Table 1 for demographic details). Studies were conducted at the Living Laboratory® at the Museum of Science, Boston. A testing station with two tables was set up in a corner of the museum exhibit. Museum patrons were free to approach researchers or were recruited by researchers and brought to the testing station as they passed through the exhibit. Prior to providing consent, participants were informed of the general research area but not of the specific question of testing the association between nonsense words and abstract visual shapes. They were briefed on general procedures, potential risks and discomforts, confidentiality of data management, and their rights as participants to withdraw at any time and their rights as participants to contact the principal investigator and the university's institutional review board if they felt the need. All adult participants provided written consent. All children participants provided written assent, and their parents or legal guardians provided written permission. After the experiment, participants were debriefed and offered stickers as a token of appreciation, as per museum protocol. All protocols were approved by the institutional review board of the University of Massachusetts Boston and the Museum of Science, Boston.

Apparatus and stimuli

Visual stimuli were presented on a laptop screen using a MacBook Pro equipped with Matlab (R2014b or 2009b) and Psychophysics Toolbox Version 3.0.12 (Brainard, 1997; Pelli, 1997). Auditory

Table 1
Participant demographics of Experiment 1 ($N = 70$).

Characteristic	*n*	Within age group:				
		% of *N*	% White	% English as (one of) first language(s)	% Spanish as (one of) first language(s)	% Bilingual
Female	39	55.71				
Age group (years)						
6–8	16	22.86	81.25	81.25	6.25	37.50
9–11	17	24.29	88.24	88.24	0	29.41
12–17	16	22.86	56.25	100.00	0	25.00
18–35	21	30	80.95	80.95	0	14.29
Unspecified	0	0.0				

stimuli were presented via noise-cancelling headphones (3 M Peltor HTB79A), and responses were recorded via a response pad (Cedrus, RB-844).

Visual shapes (~4° diameter) were created in Adobe Illustrator (see Fig. 1A). We manipulated protrusion number (three or seven), protrusion size (thin or thick), and contour spikiness (spikey or round) to create a total of four unique rounder shapes and matching spikier shapes (see Fig. 1A). Sound files—/baba/, /gaga/, /kiki/, and /titi/—were recorded using Mac OS X text-to-speech function and edited with Audacity Team (2014) to standardize sound duration and amplitude across sounds. Of note, although our auditory stimuli are generally considered nonsense words in English, the language in which the current study was conducted, it is possible that they carry meaning in other languages spoken by the participants (e.g., /kiki/ could mean bun in Spanish).

Procedure

A fixation cross (0.5° × 0.5°) appeared at screen center for trial duration (see Fig. 1B). Trials were self-initiated via button press by participants or the experimenter. To make stimulus presentation more dynamic and better engage attention, one shape loomed in (diameter increased 0–4°, 700 ms) with a "zoom" sound and remained at maximum size at screen center while two nonsense words were presented sequentially (interstimulus interval = 1 s; see Fig. 1B). Participants pressed one of two buttons to indicate which sound (first or second) best matched the given shape. Most 6- to 8-year-olds, and all older children, responded by button press. We estimate that 5 or fewer 6- to 8-year-olds (less than 31.3% of this age group) were confused by the mapping between the buttons and the response, first or second, and were given the option of saying "first" or "second," with the experimenter pressing the corresponding button. Trials were aborted and repeated if no response was made within 6 s of choice stimulus offset.

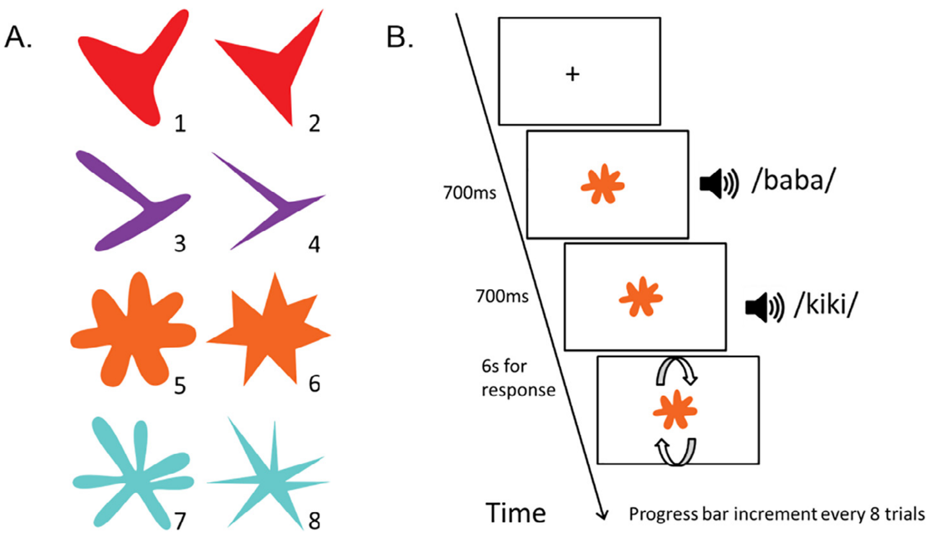

Fig. 1. Visual shapes (A) and the sequence of events in time (B) for Experiment 1. Shapes varied in protrusion number (3 protrusions: Shapes 1, 2, 3, and 4; 7 protrusions: Shapes 5, 6, 7, and 8), protrusion size, (thick: Shapes 1, 2, 5, and 6; thin: Shapes 3, 4, 7, and 8), and contour spikiness (round: Shapes 1, 3, 5, and 7; spikey: Shapes 2, 4, 6, and 8). Participants viewed an abstract shape and a pair of sounds presented sequentially (e.g., /baba/ and /kiki/). All participants were asked to indicate which sound (first or second) best matched the shape via button press. In the rare cases where the mapping between the buttons and first/second was hard to remember, especially in younger children, participants could say "first" or "second" and the experimenter would press the corresponding button. As a reward, after a choice was made, the central shape would spin (1 s at 3 Hz) and then loom out (decrease in diameter from 4° to 0°, 700 ms) along with a "zoom" sound and a progress bar incremented in length every eighth trial.

A total of 64 trials were presented, requiring roughly 10 min to complete. Two repeats of each combination (4 pairs of matched nonsense words × 8 shapes) were presented. Sound order (first or second interval) was counterbalanced across trials.

Prior to beginning the experiment, all participants completed practice trials, matching animal sounds (/woof/ or /meow/) to animal images (dog or cat), to ensure that they understood instructions. Experimental trials started once 4 consecutive practice trials were answered correctly. If this criterion was not reached within 16 practice trials, data were discarded.

Data analysis

Measuring changes in the presence of crossmodal correspondence. Choice bias in selecting which sound best matched a given shape was calculated as the proportion of trials where a /baba/ or /gaga/ sound was chosen over a /kiki/ or /titi/ sound for a given shape (*j*) by a given observer (*i*), with 0.5 subtracted (value for random choices). This is in keeping with previous research reporting similar sound-to-shape mapping responses within voiced and unvoiced obstruents (McCormick, Kim, List, & Nygaard, 2015). A choice bias of 0 indicates no bias (choice bias $_{(i,j)}$). A positive choice bias indicates an /a/ sound is preferred. A negative choice bias indicates an /i/ sound is preferred:

$$\text{Sound choice bias}_{(i,j)} = P(/\text{ba, ga/sound chosen})_{(i,j)} - 0.5. \quad (1)$$

We averaged sound choice bias across shapes for a given shape feature (e.g., contour spikiness) across participants of a given age group and compared that against zero with one-sample *t* tests. *P* values were adjusted for multiple comparison by controlling for false discovery rate using the *p.adjust* function in R (R Development Core Team, 2011). A significant result indicates a significant association in the corresponding age group for the shape feature of interest. Hedges' *g*, equivalent to Cohen's *d* adjusted for small sample size (Lakens, 2013; Olejnik & Algina, 2003), was estimated by the R effsize package *cohen.d* function. We reported the absolute Hedges' *g* to remove the sign of the choice bias for easier comparison across shape features.

Measuring changes in the magnitude of crossmodal correspondence. To explore whether and how the magnitude of crossmodal association of each shape feature changes across development, mixed effect logistic regression modeling was performed on raw data using the R lme4 package *glmer* function. The dependent variable of the analysis was a binary outcome of whether the /baba/ or /gaga/ sound was chosen over /kiki/ or /titi/ for each trial. We built our null model based on previous literature (Chen et al., 2016), which showed that three shape features are significant predictors of adult participants' sound choice. The null model included only three categorical fixed factors: shape features (contour spikiness, protrusion size, and protrusion number) and subject as random effects contributing to the intercept of the model as predictors. In the alternative model, in addition to the predictors included in the null model, age group and its interaction with each feature were also included as predictors. To simplify comparisons, we included only data from younger children (6- to 8-year-olds) and adults (18- to 35-year-olds), so age group was a binary categorical predictor. If the magnitude of crossmodal correspondence changes across development, we expected that including age group as an interaction term with shape features would improve model fit.

Results

Data reflect measures from 70 participants. Data from an additional 3 participants who completed fewer than 32 trials were excluded. All participants reached the practice criterion within 20 practice trials, with only 1 participant requiring the maximum number of practice trials. Participants included for analysis completed 4 (minimum) to 20 (maximum) practice trials before reaching the criterion ($M = 5.24$, $SD = 2.46$). Of a maximum of 64 possible trials, an average of 61.6 trials were completed (see Fig. 2A for average trial numbers per age group).

We collapsed across nonsignificant factors: gender ($F\text{s} \leq 1.089$, $p\text{s} \geq .301$, $\eta^2_G\text{s} \leq 0.011$, $\omega^2\text{s} \leq 0.0019$), temporal intervals (whether the /a/ related sound was presented first or second: $F\text{s} \leq 3.345$, $p\text{s} \geq .072$, $\eta^2_G\text{s} \leq 0.007$, $\omega^2\text{s} \leq 0.0314$), and consonants (whether /baba/ vs. /gaga/ or /kiki/

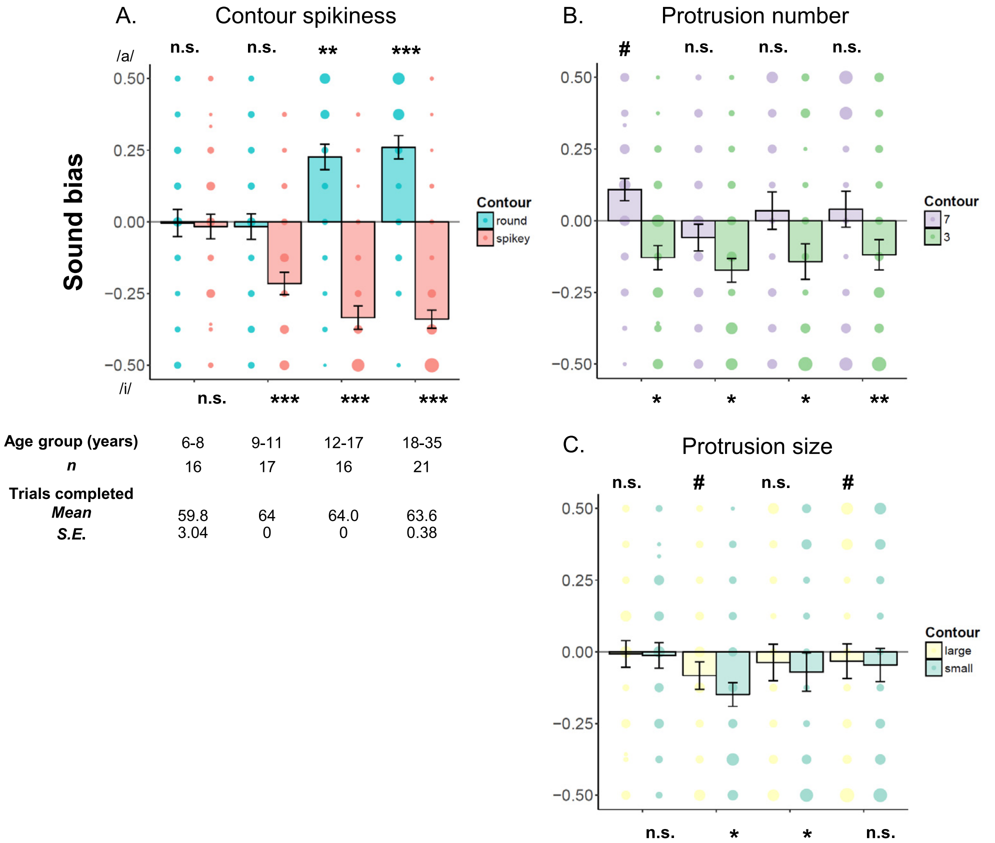

Fig. 2. Mean choice bias results revealing effects of contour spikiness (A), protrusion number (B), and protrusion size (C) in Experiment 1, with error bars indicating standard error of the mean across participants and dots showing individual data, with dot size scaled by number of participants with data at the same value. *$p < .05$, **$p < .01$, ***$p < .001$; #$p < 0.1$; n.s., not significant.

vs. /titi/ was presented: $Fs \leq 0.148$, $ps \geq .701$, $\eta^2_G s \leq 0.001$, $\omega^2 s \leq -0.0133$). Because in each age group a majority of participants ($\geq$80.95%) reported English as their first language, or one of their first languages if they were bilingual, with sparse contribution from other languages, we included all participants in the subsequent analysis. The results were similar when bilingual participants (see Supplementary Fig. 1A–C in online supplementary material) or Spanish-speaking participants (Supplementary Fig. 1E–G) were excluded.

Developmental changes in the presence/absence of crossmodal correspondences

Fig. 2 plots mean sound choice bias as a function of arbitrary age bins for contour spikiness (Fig. 2A), protrusion number (Fig. 2B), and protrusion size (Fig. 2C). Data were collapsed across similar shape features for contour spikiness (e.g., round shapes (see Fig. 1A): Shapes 1, 3, 5, and 7; spikey shapes (see Fig. 1A): Shapes 2, 4, 6, and 8), for protrusion number (e.g., 3 protrusions (see Fig. 1A): Shapes 1, 2, 3, and 4; 7 protrusions (see Fig. 1A): Shapes 5, 6, 7, and 8), and for protrusion size (e.g., thick (see Fig. 1A): Shapes 1, 3, 5, and 7; thin (see Fig. 1A): Shapes 2, 4, 6, and 8). A positive value indicates stronger preferences for correspondences with /a/ sounds, whereas a negative value indicates stronger preferences for /i/ sounds.

We did not find a significant association between nonsense words and contour spikiness in the youngest age group (6- to 8-year-olds); mean choice bias for contour was not significantly different from zero ($ps \geq .589$, $gs \leq 0.135$). Mean choice bias became significantly different from zero with age, with bigger divergence between spikey and round shapes. For example, we observed a significant association in the expected direction (round shapes with /a/ sounds and spikey shapes with /i/ sounds) in both older children (12- to 17-year-olds) and adults (18- to 35-year-olds) ($ps \leq .0012$, $gs \geq 0.972$). In 9- to 11-year-olds, a spikey shape was significantly associated with an /i/ sound, $t(16) = -4.4806$, $p < .001$, $g = 1.878$, but a round shape was not significantly associated with an /a/ sound, $t(16) = -0.30266$, $p = .766$, $g = -0.072$. This suggests that the association between spikey shapes and /i/ sounds is relatively more robust and comes online earlier than the association between round shapes and /a/ sounds.

Our analysis of the association between nonsense words and protrusion number and size yielded three findings of interest. First, across age groups, there was a consistent association between 3-protrusion shapes and /i/ sounds ($ps \leq .026$, $gs \geq 0.686$). Second, 6- to 8-year-olds marginally associated 7-protrusion shapes with /a/ sounds, $t(15) = 2.2144$, $p = .068$, $g = 0.540$, a unique pattern not observed in other age groups ($ps \geq .222$, $|g|s \leq 0.342$). The marginal association between 7-protrusion shapes with /a/ sounds was weakened when data from bilingual or Spanish-speaking participants were excluded from analysis (see Supplementary Fig. 1B). Third, 9- to 11-year-olds and 12- to 17-year-olds associated small-protrusion shapes with /i/ sounds ($ps < .018$, $gs \geq 0.814$), with no significant associations for large-protrusion shapes ($ps \geq .081$, $g = 0.276$), whereas 6- to 8-year-olds and adults showed no significant associations for protrusion size ($ps \geq .083$, $gs \leq 0.467$).

Developmental changes in the magnitude of crossmodal correspondences

To quantify changes in the magnitude of crossmodal correspondences, we performed a mixed effect binary/logistic regression to predict participants' odds of choosing a /baba/ or /gaga/ sound as opposed to a /kiki/ or /titi/ sound when a given shape feature was presented. Two models were compared against each other, with both including three types of shape features as predictors based on previous literature (Chen et al., 2016). The alternative model (Akaike's information criterion [AIC] = 2530.8, Bayesian information criterion [BIC] = 2582.5, log likelihood = −1256.4), which included age group and interaction terms with age group, showed significantly improved model fit, $\chi^2(4) = 238.11$, $p < .001$, compared with the null model (AIC = 2760.9, BIC = 2789.6, log likelihood = −1375.5), which did not include age group and interaction terms with age group. The logit coefficients of each predictor in the alternative model are reported in Table 2. To interpret the coefficients, we converted the exponential of the coefficients into odds. Changing the contour spikiness of a shape alone did not change the odds of choosing an /a/ sound over an /i/ sound ($p = .663$). However, there was a significant interaction between contour spikiness and age group; the odds of choosing an /a/ sound increased by a factor of 20.22 ($p < .001$) when the shape was round as opposed to spikey. This was true in adults but not in 6- to 8-year-olds, suggesting that the influence of contour spikiness on cross-modal correspondence is not stable across development. On the other hand, we did not observe an interaction between age group and protrusion number ($p = .547$); increasing the protrusion number from 3 to 7 increased the odds of choosing an /a/ sound by a factor of 2.64 independent of age group ($p < .001$). We speculate that this may be driven by the robust association between 3-protrusion shapes and /i/ sounds found across age groups, which may mask the association between 7-protrusion shapes and /a/ sounds seen in the graphs for 6- to 8-year-olds.

Discussion

Experiment 1 showed that 6- to 8-year-olds did not associate /a/ sounds with round shapes and /i/ sounds with spikey shapes, unlike older children (12- to 17-year-olds) and adults who showed the expected associations. Rather, 6- to 8-year-olds tend to prioritize protrusion number, associating 7-protrusion shapes, irrespective of shape contour, with /a/ sounds, an association not seen in any other age group and a result further supported by a significant interaction between age group and contour spikiness in the mixed effect logistic regression. This developmental change gives rise to two speculations. First, the manipulation of contour spikiness in our shapes might not be sufficient to induce sound–shape correspondence in 6- to 8-year-olds. Second, there is a qualitative change in the feature prioritized for sound–shape correspondence in 6- to 8-year-olds compared with older children,

Table 2
Coefficients of the final mixed effect logistic regression model predicting sound bias.

Fixed effects	Odds ratio	*SE*	*z* Value	*p* Value	95% Confidence interval	
(Intercept)	0.533	0.1730	−3.641	.001	[0.319,	0.748]
Age group 0: 6–8 years 1: 18–35 years	0.171	0.2570	−6.866	<.00	[0.104,	0.283]
Contour spikiness 0: spiky 1: rounded	1.061	0.1356	0.436	.663	[0.813,	1.384]
Protrusion number 0: 3 protrusions 1: 7 protrusions	2.640	0.1362	7.128	<.001	[2.022,	3.448]
Protrusion size 0: small 1: large	1.157	0.1357	1.076	.282	[0.887,	1.510]
Age group × contour spikiness	20.222	0.2057	14.619	<.001	[13.513,	30.262]
Age group × protrusion number	1.129	0.2021	0.602	.547	[0.760,	1.678]
Age group × protrusion size	0.946	0.1966	−0.281	.779	[0.644,	1.391]

Note. Number of observations: 2292, grouped by subject ID ($n = 37$). Random effects: subject (intercept) variance = 0.1757, $SD = 0.4192$.

especially in situations like the current experimental paradigm where children were free to select the shape feature they used as the basis for their judgments.

To tease apart these possibilities, we highlighted contour spikiness in Experiment 2 by presenting two shapes having the same number and size of protrusions, one round and one spikey, side by side for direct comparison. Participants selected which of the two shapes best matched a given sound. If the absence of an association between contour spikiness and sound in young children in Experiment 1 was because our manipulation of contour spikiness was not robust enough, then we expected that 6- to 8-year-olds would not exhibit the expected association even when the relevant feature was highlighted in Experiment 2. However, if the absence of an association was because 6- to 8-year-olds prioritized different features of the visual shape, then 6- to 8-year-olds would exhibit the expected association when the relevant feature was highlighted in Experiment 2.

Experiment 2: Two-alternative forced choice (shape bias with contour spikiness highlighted)

Method

Participants

A total of 21 adults (18–35 years old) and 47 children (6–17 years old) participated in Experiment 2 (see Table 3 for demographic details).

Apparatus and stimuli

The apparatus and stimuli were identical to those of Experiment 1 except that shapes were presented simultaneously side by side. There were four unique rounder shapes matched to spikier shapes (see Fig. 1A for matching shape pairs: 1–2, 3–4, 5–6, and 7–8).

Procedure

A fixation cross ($0.5° \times 0.5°$) appeared at screen center for trial duration (see Fig. 3). Trials were self-initiated via button press by participants or the experimenter. Two shapes, one round and one spikey, matched along other dimensions, loomed in (diameter increased 0–4°, 700 ms) with a "zoom" sound and remained at maximum size, left and right of screen center (3.19°), as a nonsense word was presented. Participants indicated the side (left or right) with the shape best matching the sound via button press. Most 6- to 8-year-olds, and all older children, responded by button press. We

Table 3
Participant demographics of Experiment 2 ($N = 68$).

Characteristic	n	Within age group: % of N	% White	% English as (one of) first language(s)	% Spanish as (one of) first language(s)	% Bilingual
Female	33	48.53				
Age group (years)						
6–8	15	22.06	53.33	100.00	6.67	13.33
9–11	18	26.47	50.00	88.89	5.56	16.67
12–17	14	20.59	64.29	100.00	0.00	21.43
18–35	21	30.88	61.90	95.24	9.52	28.57

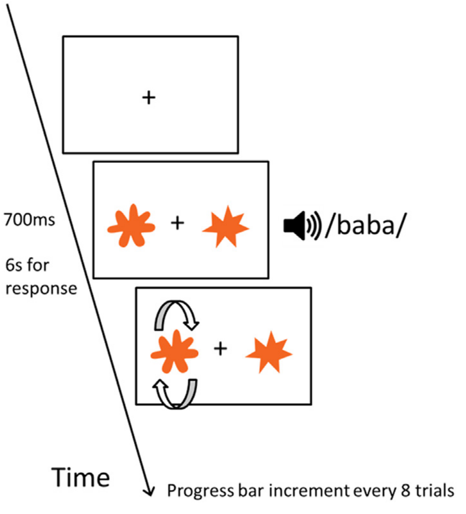

Fig. 3. Procedure of Experiment 2. A pair of shapes were presented simultaneously along with a nonsense word (/baba/, /gaga/, /kiki/, or /titi/). All participants were asked to indicate which shape best matched the sound via button press (left or right). In the rare cases where younger participants had trouble mapping the location of the shapes to the buttons, they were given the option to respond by pointing. As a reward, after a choice was made, the chosen shape would spin (1 s at 3 Hz) and then loom out (decrease in diameter from 4° to 0°, 700 ms) along with a "zoom" sound and a progress bar incremented in length every eighth trial.

estimate that 2 or fewer 6- to 8-year-olds (less than 13.3% of this age group) were confused by the mapping between the buttons and left/right and were given the option to respond by pointing, with the experimenter pressing the corresponding button. Trials were aborted and repeated if no response was made within 6 s of choice stimulus offset. As in Experiment 1, a total of 64 trials were presented, including 4 repeats of each combination (4 pairs of matched shapes × 4 nonsense words), with shape location counterbalanced across trials. As before, all participants completed practice trials to ensure that they understood instructions.

Analysis

As in Experiment 1, association strength was quantified for each individual by measuring choice bias except that the choice bias was calculated as the proportion of trials where a round shape was chosen over a spikey shape for a given sound (j) by a given observer (i), with 0.5 subtracted (value for random choices). A choice bias of 0 indicates no bias (choice bias $_{(i,j)}$). A positive choice bias indicates a round shape is preferred. A negative choice bias indicates a spikey shape is preferred:

$$\text{Shape choice bias}_{(i,j)} = P(\text{round shape chosen})_{(i,j)} - 0.5. \quad (2)$$

We averaged shape choice bias across sounds for a given vowel (e.g., across participants of a given age group) and compared that against zero with one-sample *t* tests. We conducted mixed effects logistic regression modeling as in Experiment 1 except that here the dependent variable was a binary outcome of whether a round shape was chosen over a spikey shape. We built our null model based on previous literature (Maurer et al., 2006), which showed that nonsense sound categories (e.g., /baba/ vs. /kiki/) modulate participants' shape choice. In the null model, we included a categorical fixed factor, sound category (/ba, ga/ vs. /ki, ti/) and subject as random effects contributing to the intercept of the model as predictors. In the alternative model, in addition to the predictors included in the null model, we also included age group and its interaction with sound category as predictors. To simplify comparisons, we included only data from younger children (6- to 8-year-olds) and adults (18- to 35-year-olds). If the odds of choosing a round shape for a given sound category change with age, we expected to see improved fit for the alternative model compared with the null model.

Results

Data reflect measures from 68 participants. Data from an additional 2 participants who completed fewer than 16 trials were excluded. All participants reached the criterion within 20 practice trials, with no participant requiring the maximum number of practice trials. Participants included for analysis completed 4 (minimum) to 13 practice trials before reaching the practice criterion ($M = 5.24$, $SD = 2.46$). Of a maximum of 64 possible trials, an average of 61.6 trials were completed (see Fig. 4 for average trial numbers per age group). Because in each age group a majority of participants ($\geq$88.89%) reported English as their first language, or as one of their first languages if they were bilingual, with sparse contributions from other languages, we included all participants in the subsequent analysis. We found similar results when bilingual participants (see Supplementary Fig. 1D) or Spanish-speaking participants (see Supplementary Fig. 1H) were removed.

We collapsed across nonsignificant factors: gender (Fs ≤ 1.170, ps $\geq .284$, η_G^2s ≤ 0.013, ω^2s ≤ 0.001) and spatial location (whether the round shape was presented left or right) (Fs ≤ 1.056, ps $\geq .308$, η_G^2s ≤ 0.0008, ω^2s ≤ 0.0008).

Robust presence of crossmodal correspondences when contour spikiness is highlighted

Fig. 4 plots mean shape choice bias as a function of arbitrary age bins for vowel sounds. Data were collapsed across vowel sounds (e.g., /a/ sounds included /baba/ and /gaga/, /i/ sounds included /kiki/ and /titi/). A positive value indicates stronger preferences for round shapes, whereas a negative value indicates stronger preferences for spikey shapes. Children and adults of all age groups exhibited the expected association, matching round shapes with /a/ sounds and spikey shapes with /i/ sounds (ps $\leq .0405$, gs ≥ 0.567). Thus, when contour spikiness was highlighted, our particular shape stimuli were sufficient to solicit significant correspondences between sounds and contour spikiness in 6- to 8-year-olds, rejecting the speculation that our manipulation of contour spikiness was not robust enough for young children.

Developmental changes in the magnitude of crossmodal correspondences

Despite the significant association found in younger age groups, the strength of crossmodal correspondence may change across development. To quantify changes in the magnitude of crossmodal correspondences, we performed a mixed effect binary/logistic regression to predict participants' odds of choosing a round shape as opposed to a spikey shape when a given sound category was presented. The alternative model (AIC = 2411.5, BIC = 2440.1, log likelihood = −1200.7), including age group and interaction terms with age group, significantly improved model fit, $\chi^2(2) = 24.493$, $p < .001$, compared with the null model (AIC = 2432.0.8, BIC = 2449.1, log likelihood = −1213.0). The logit coefficients of each predictor in the alternative model are reported in Table 4. To interpret the coefficients, we took the exponential of the coefficients to convert them into odds. Changing the sound category from an /i/ sound (e.g., /kiki/, /titi) to an /a/ sound (e.g., /baba/, /gaga/) increased the odds of participants choosing round shapes over spikey shapes by a factor of 6.231 ($p < .001$) regardless of age. In addition, the odds

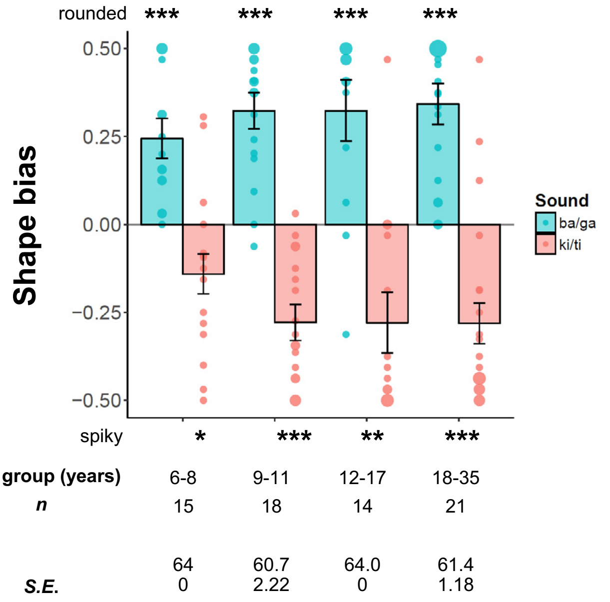

Fig. 4. Mean choice bias results revealing effects of sound category in Experiment 2, with error bars indicating standard error of the mean across participants and dots showing individual data, with dot size scaled by number of participants with data at the same value. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Table 4
Coefficients of the final mixed effect logistic regression model predicting shape bias.

Fixed effects	Odds ratio	*SE*	*z* Value	*p* Value	95% Confidence interval
(Intercept)	0.527	0.2415	−2.653	.008	[0.328, 0.846]
Age group	0.567	0.3179	−1.788	.0739	[0.304, 1.056]
0: 6–8 years					
1: 18–35 years					
Sound category	6.231	0.1529	11.966	<.001	[4.617, 8.408]
0: /kiki/ or /titi/					
1: /baba/ or /gaga/					
Age group × sound category	2.871	0.2136	4.937	<.001	[1.889, 4.364]

Note. Number of observation: 2277, grouped by subject ID ($n = 37$). Random effects: subject (intercept) variance = 0.7194, *SD* = 0.8482.

of choosing a round shape increased by a factor of 2.871 ($p < .001$), for /a/ sounds compared with /i/ sounds in adults but not in 6- to 8-year-olds. Importantly, this is not because adults are choosing more round shapes in general given that the odds for adults choosing a round shape, regardless of sound category, marginally decreased by a factor of 0.567 ($p = .074$) compared with children. Rather, even when the relevant feature was highlighted, adults still exhibited stronger sound–shape correspondence by contour spikiness than 6- to 8-year-olds.

General discussion

Our study seeks to understand the qualitative and quantitative changes in sound–shape cross-modal correspondence across development and the factors underlying these changes. In Experiment 1, when the relevant shape feature was not highlighted and participants were free to choose the feature on which to base their judgments, young children (6- to 8-year-olds) made associations based on protrusion number but not contour spikiness, unlike adults, who made associations based on contour spikiness. This suggests a *qualitative change* in the feature prioritized for sound–shape correspondence in younger children compared with adults. In Experiment 2, when shapes were side by side, allowing a direct comparison and highlighting shape contour spikiness as the relevant feature, young children (6- to 8-year-olds) resembled adults and showed the expected associations based on contour spikiness, associating round shapes with /a/ sounds and spikey shapes with /i/ sounds. This suggests that visual contextual comparison modulates the exhibition of sound–shape correspondence. In addition, whereas both young children and adults exhibited significant associations in the expected direction, the magnitude of the effect was stronger in adults than in young children, suggesting a *quantitative change* in the strength of crossmodal correspondence across development.

Potential reasons for inconsistencies with previous developmental reports

We found qualitative and quantitative changes in crossmodal correspondence across development. Why do our results from Experiment 1 seem inconsistent with previous findings of adult-like sound–shape correspondence in younger children (Asano et al., 2015; Maurer et al., 2006; Ozturk et al., 2013; Pejovic & Molnar, 2017; Spector & Maurer, 2013; Tzeng et al., 2017)? It is unlikely that younger children had trouble understanding instructions in our task given that we excluded participants who could not reach the criterion on practice trials. Nor is it likely that younger children had trouble attending to our task given that they completed approximately the same number of trials as older children and adults. Furthermore, it is unlikely that our use of a forced-choice task, requiring explicit responses, was too demanding for young children. Previous work shows that implicit measures (looking time: Ozturk et al., 2013; yes/no matching task: Nava, Grassi, & Turati, 2016; electroencephalography: Asano et al., 2015), as well as explicit measures (forced choice: Maurer et al., 2006; Spector & Maurer, 2013; rating task: Marks, Hammeal, & Bornstein, 1987), can show a bouba-kiki effect in younger children. Rather, participants of all ages exhibited significant sound–shape correspondence based on at least one shape feature and, thus, were performing our task but prioritizing different features.

Several factors can explain why our findings differ from previous findings. One factor is the role of context. For example, although Maurer et al. (2006) and Tzeng et al. (2017) tested young children's explicit choice, as did our study, they presented two shapes and two labels, providing visual contrast as well as auditory contrast, unlike our study, which provided either auditory contrast (Experiment 1) or visual contrast (Experiment 2) but not both at the same time. It is possible that sensitivity to, or reliance on, context depends on the implicit/explicit nature of the task because the availability of contrast/context seems less relevant in studies using implicit measures (e.g., Asano et al., 2015; Ozturk et al., 2013), where one shape and one label are often presented without contrast. A second factor is the use of additional information by young children. For example, Maurer et al. (2006) and Spector and Maurer (2013) had the experimenter pronounce the nonsense words in front of the children. Therefore, movements of the mouth in producing the sound were available to children, whereas in our study prerecorded nonsense words were presented purely aurally. Finally, whether and how the relevant feature might be highlighted needs to be considered. For example, Tzeng et al. (2017) had children point to the round versus spikey shapes before testing for sound–shape correspondence, which can intentionally or unintentionally highlight the relevant feature for children to make successful correspondence, similar to our Experiment 2.

Speculated factors underlying the developmental changes

Below we speculate on how qualitative and quantitative changes in sound–shape correspondence from young childhood to adulthood reveal intrinsic limitations on how children might select and process visual shape features differently from adults.

First, qualitative change might arise from children adopting different perceptual processing styles than adults. In Experiment 1, 6- to 8-year-olds showed strong associations based on protrusion number but not contour spikiness, whereas adults showed strong associations based on contour spikiness and only weak associations based on protrusion number. According to our conceptualization, this would imply a switch from prioritizing a global feature (protrusion number) to prioritizing a local feature (contour spikiness) across development. A switch in perceptual processing style from global to local processing has also been found in other developmental studies (e.g., Kimchi et al., 2005; Kovshoff et al., 2015). Our result provides further support that sound–shape correspondence might originate from mid-level interactions between abstract shapes and sounds being modulated by perceptual processing style (Chen et al., 2016). However, our result does not speak to the development of perceptual grouping because it cannot be compared with traditional studies testing perceptual processing style owing to differences in stimuli, (e.g., our shapes are not composed of individual local elements like Navon patterns) and differences in procedure (e.g., our paradigm is not a speeded judgment task).

Second, quantitative change might arise from children being less efficient in extracting relevant shape features based on visual context than adults. When contour spikiness was highlighted in Experiment 2, 6- to 8-year-olds showed significant associations between contour spikiness and sounds, an association not observed when the relevant feature was not highlighted in Experiment 1. This suggests that children are sensitive to visual context and use visual context to extract relevant information (e.g., Waxman & Klibanoff, 2000). Other studies also find that crossmodal correspondences are susceptible to contextual influences (e.g., pitch and height, brightness and size) and may vary based on the relative mapping of stimulus features (e.g., the higher of two tones, the spikier of two shapes) (Chiou & Rich, 2012; Walker 2012; Walker, Walker, & Francis, 2015). For example, the pitch of a sound modulates orienting toward a spatial location based on pitch–height correspondence (Chiou & Rich, 2012). Thus, the orienting influence of a sound depends on context; the same sound can direct attention to the lower visual field, when presented with another higher pitch sound as comparison, or to the higher visual field, when presented with a lower pitch sound. Thus, pitch–height correspondence is based not on absolute pitch but rather on relative pitch. Our study extends such findings for sound–shape correspondence, showing that making relative contrast explicit might be necessary for younger children to exhibit the same qualitative associations seen in adults.

Despite being sensitive to visual context, younger children may also be less efficient than adults in extracting the relevant information, which may give rise to quantitative differences. Thus, younger children might require more contrast between shape pairs to extract relevant features (see Smith commentary in Marks et al., 1987) or might require more exemplars to extract or retain the relevant information (Casasola & Park, 2013; Twomey, Ranson, & Horst, 2014; cf. Maguire, Hirsh-Pasek, Golinkoff, & Brandone, 2008). As a result, compared with adults, young children exhibited significant but weaker associations in Experiment 2, as indicated by the significant interaction of age and sound category in the logistic regression model.

Broader implications

Beyond the development of sound–shape correspondence, our study also highlights the distinct impact of three shape features on sound–shape correspondences. For vowel–size correspondences, /a/ sounds are associated with larger surface area than /i/ sounds (e.g., Ohtake & Haryu, 2013; Peña, Mehler, & Nespor, 2011; Sapir, 1929). All three of the features we tested can be related to surface area. More specifically, protrusion number reflects total surface area across protrusions, being larger for more protrusions; protrusion size reflects surface area of individual lobes, being larger for larger protrusions; and contour spikiness reflects surface area at the tip of the protrusion, being larger for rounder protrusions. Nevertheless, we see distinct associations based on the feature, with protrusion number being prioritized in younger children and contour spikiness being prioritized in older children and adults. Of note, the null effects seen for protrusion size suggest that participants did not merely base judgments on surface area; instead they based judgments on particular features. Our work finds unique contributions for different visual features, extending the findings of Chen et al. (2016) with a novel set of shape stimuli and a developmental sample.

Despite supporting the overall conclusion of Chen et al. (2016), we found opposite effects of increasing protrusion number. In our study, as protrusion number increased in Experiment 1, adults were more likely to match the shape with /a/-related sounds, whereas in Chen and colleagues' study adults were more likely to match the shape with /i/-related sounds. We speculate that this difference is driven by differences in how surface area was controlled in the two studies. As protrusion number increased, the surface area of individual protrusions decreased in Chen and colleagues' study, wheres the surface area of individual protrusions was kept constant, which resulted in larger overall shape surface area in our study. Of note, children as young as 3 years have been found to prioritize number over other dimensions of quantity such as overall surface area (Cantlon et al., 2010). Our study did not specifically test the relative salience of number of protrusions versus overall shape surface area. Although 6- to 8-year-olds in our study prioritized number when no visual context was provided (Experiment1), increases in protrusion number were confounded with increases in overall surface area. Furthermore, it is not clear whether protrusion number would have been prioritized if the two features, protrusion number and surface area, had been compared directly using a two-alternative forced-choice design as used for Experiment 2, where protrusion number was kept constant when contrasting stimulus pairs. Overall, little is known regarding the relative salience of different shape features, and future work is needed to dissociate the influence of number of protrusions, total surface area across protrusions, and surface area of individual protrusions on sound–shape correspondence.

In summary, our study is the first to compare how unique features of a shape might contribute to the bouba-kiki effect differently in children and adults. We found that younger children prioritize number of protrusion over contour spikiness as compared with older children and adults. Our results highlight the contribution of changes in perceptual processing style and changes in the influence of context on shape processing as sound–shape correspondence develops.

Acknowledgments

The authors thank a valuable team of undergraduate research apprentices who helped with data collection. We also thank Zsuzsa Kaldy, Ed Tronick, David Lewkowicz, Adam Jacobs, Alice Carter, and three anonymous reviewers for helpful comments on earlier drafts of this manuscript. Finally, none of this work would have been possible without the support of the staff of the Living Laboratory®, Museum of Science, Boston, where we conducted our studies, nor the generous support of the UMB Dean's office (vmc).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jecp.2018.10.005.

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.

Belli, F., & Sagrillo, F. (2001). *Qual è Takete? Qual è Maluma? La psicolinguistica applicata alla comunicazione pubblicitaria [What is Takete? What is Maluma? Psycholinguistics applied to advertising]* (2nd ed.). Milan, Italy: Franco Angeli.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433–436.

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*, 165–172.

Cantlon, J. F., Safford, K. E., & Brannon, E. M. (2010). Spontaneous analog number representations in 3-year-old children. *Developmental Science, 13*(2), 289–297.

Casasola, M., & Park, Y. (2013). Developmental changes in infant spatial categorization: When more is best and when less is enough. *Child Development, 84*, 1004–1019.

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

Chiou, R., & Rich, A. N. (2012). Cross-modality correspondence between pitch and spatial location modulates attentional orienting. *Perception, 41*, 339–353.

Davis, R. (1961). The fitness of names to drawings: A cross-cultural study in Tanganyika. *British Journal of Psychology, 52*, 259–268.

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

Hamilton-Fletcher, G., Wright, T. D., & Ward, J. (2015). Cross-modal correspondences enhance performance on a colour-to-sound substitution device. *Multisensory Research, 29*, 337–363.
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*. https://doi.org/10.1098/rstb.2013.0298.
Imai, M., Miyazaki, M., Yeung, H. H., Hidaka, S., Kantartzis, K., Okada, H., & Kita, S. (2015). Sound symbolism facilitates word learning in 14-month-olds. *PLoS One, 10*(2), e116494.
Irwin, F. W., & Newland, E. (1940). A genetic study of the naming of visual figures. *Journal of Psychology, 9*(1), 3–16.
Kimchi, R., Hadad, B., Behrmann, M., & Palmer, S. E. (2005). Microgenesis and ontogenesis of perceptual organization: Evidence from global and local processing of hierarchical patterns. *Psychological Science, 16*, 282–290.
Köhler, W. (1929). *Gestalt psychology*. New York: Liveright.
Köhler, W. (1947). *Gestalt psychology* (2nd ed.). New York: Liveright.
Kovshoff, H., Iarocci, G., Shore, D. I., & Burack, J. A. (2015). Developmental trajectories of form perception: A story of attention. *Developmental Psychology, 51*, 1544–1552.
Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for *t*-tests and ANOVAs. *Frontiers in Psychology, 4*. https://doi.org/10.3389/fpsyg.2013.00863.
Lewkowicz, D. J., & Minar, N. J. (2014). Infants are not sensitive to synaesthetic cross-modality correspondences: A comment on Walker et al. (2010). *Psychological Science, 25*, 832–834.
Lupyan, G., & Casasanto, D. (2015). Meaningless words promote meaningful categorization. *Language and Cognition, 7*, 167–193.
Maguire, M. J., Hirsh-Pasek, K., Golinkoff, R. M., & Brandone, A. C. (2008). Focusing on the relation: Fewer exemplars facilitate children's initial verb learning and extension. *Developmental Science, 11*, 628–634.
Marks, L. E., Hammeal, R. J., & Bornstein, M. H. (1987). Perceiving similarity and comprehending metaphor. With commentary by Linda B. Smith. *Monographs of Society of Research in Child Development, 52* (1, Serial No. 215).
Maurer, D., & Mondloch, C. (2004). Neonatal synesthesia: A reevaluation. In L. Robertson & N. Sagiv (Eds.), *Attention on synesthesia: Cognition, development and neuroscience* (pp. 193–213). Oxford, UK: Oxford University Press.
Maurer, D., Pathman, T., & Mondloch, C. J. (2006). The shape of boubas: Sound–shape correspondences in toddlers and adults. *Developmental Science, 9*, 316–322.
McCormick, K., Kim, J. Y., List, S., & Nygaard, L. C. (2015). Sound to meaning mappings in the bouba-kiki effect. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.), *Proceedings of the 37th annual conference of the Cognitive Science Society* (pp. 1565–1580). Austin, TX: Cognitive Science Society.
Nava, E., Grassi, M., & Turati, C. (2016). Audio-visual, visuo-tactile and audio-tactile correspondences in preschoolers. *Multisensory Research, 29*, 93–111.
Nayar, K., Franchak, J., Adolph, K., & Kiorpes, L. (2015). From local to global processing: The development of illusory contour perception. *Journal of Experimental Child Psychology, 131*, 38–55.
Nielsen, A., & Rendall, D. (2011). The sound of round: Evaluating the sound-symbolic role of consonants in the classic Takete-Maluma phenomenon. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale, 65*, 115–124.
Nielsen, A. K. S., & Rendall, D. (2013). Parsing the role of consonants versus vowels in the classic Takete-Maluma phenomenon. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale, 67*, 153–163.
Ohtake, Y., & Haryu, E. (2013). Investigation of the process underpinning vowel–size correspondence. *Japanese Psychological Research, 55*(4). https://doi.org/10.1111/jpr.12029.
Oishi, S., Jaswal, V. K., Lillard, A. S., Mizokawa, A., Hitokoto, H., & Tsutsui, Y. (2014). Cultural variations in global versus local processing: A developmental perspective. *Developmental Psychology, 50*, 2654–2665.
Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: Measures of effect size for some common research designs. *Psychological Methods, 8*, 434–447.
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*, 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*, 581–586.
Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437–442.
Peña, M., Mehler, J., & Nespor, M. (2011). The role of audiovisual processing in early conceptual development. *Psychological Science, 22*, 1419–1421.
Poirel, N., Mellet, E., Houdé, O., & Pineau, A. (2008). First came the trees, then the forest: Developmental changes during childhood in the processing of visual local–global patterns according to the meaningfulness of the stimuli. *Developmental Psychology, 44*, 245–253.
Poirel, N., Simon, G., Cassotti, M., Leroux, G., Perchey, G., Lanoë, C., ... Houdé, O. (2011). The shift from local to global visual processing in 6-year-old children is associated with grey matter loss. *PLoS One, 6*(6), e20879.
Puspitawati, I., Jebrane, A., & Vinter, A. (2014). Local and global processing in blind and sighted children in a naming and drawing task. *Child Development, 85*, 1077–1090.
R Development Core Team (2011). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
Ramachandran, V. S., & Hubbard, E. M. (2003). Hearing colors, tasting shapes. *Scientific American*, 52–59.
Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology, 12*, 225–239.
Scherf, K. S., Behrmann, M., Kimchi, R., & Luna, B. (2009). Emergence of global shape processing continues through adolescence. *Child Development, 80*, 162–177.
Spector, F., & Maurer, D. (2013). Early sound symbolism for vowel sounds. *i-Perception, 4*, 239–241.
Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics, 73*, 971–995.
Spence, C., & Parise, C. V. (2012). The cognitive neuroscience of crossmodal correspondences. *i-Perception, 3*, 410–412.
Stiles, N. R. B., & Shimojo, S. (2015). Auditory sensory substitution is intuitive and automatic with texture stimuli. *Scientific Reports, 5*, 15628.

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. https://doi.org/10.1177/2041669517724807.
Sweeny, T. D., Guzman-Martinez, E., Ortega, L., Grabowecky, M., & Suzuki, S. (2012). Sounds exaggerate visual shape. *Cognition, 124*, 194–200.
Audacity Team (2014). Audacity: Free audio editor and recorder [computer program]. Version 2.0.0 retrieved April 20, 2014, from http://audacity.sourceforge.net.
Twomey, K. E., Ranson, S. L., & Horst, J. S. (2014). That's more like it: Multiple exemplars facilitate word learning. *Infant and Child Development, 23*, 105–122.
Tzeng, C. Y., Nygaard, L. C., & Namy, L. L. (2017). Developmental change in children's sensitivity to sound symbolism. *Journal of Experimental Child Psychology, 160*, 107–118.
Walker, P. (2012). Cross-sensory correspondences and cross talk between dimensions of connotative meaning: Visual angularity is hard, high-pitched, and bright. *Attention, Perception, & Psychophysics, 74*, 1792–1809.
Walker, P., Bremner, J. G., Mason, U., Spring, J., Mattock, K., Slater, A., & Johnson, S. P. (2010). Preverbal infants' sensitivity to synaesthetic cross-modality correspondences. *Psychological Science, 21*, 21–25.
Walker, P., Bremner, J. G., Mason, U., Spring, J., Mattock, K., Slater, A., & Johnson, S. P. (2014). Preverbal infants are sensitive to crossmodal correspondences: Much ado about the null results of Lewkowicz and Minar (2014). *Psychological Science, 25*, 835–836.
Walker, P., Walker, L., & Francis, B. (2015). The size–brightness correspondence: Evidence for crosstalk among aligned conceptual feature dimensions. *Attention, Perception, & Psychophysics, 77*, 2694–2710.
Waxman, S. R., & Klibanoff, R. S. (2000). The role of comparison in the extension of novel adjectives. *Developmental Psychology, 36*, 571–581.
Westbury, C. (2005). Implicit sound symbolism in lexical access: Evidence from an interference task. *Brain and Language, 93*, 10–19.