Journal of Experimental Child Psychology 209 (2021) 105167

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

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

The *feeling* of “kiki”: Comparing developmental changes in sound–shape correspondence for audio–visual and audio–tactile stimuli

Hiu Mei Chow [a,b], Daniel A. Harris [a,c], Sandy Eid [a], Vivian M. Ciaramitaro [a,*]

[a] *Department of Psychology, Developmental and Brain Sciences, University of Massachusetts Boston, Boston, MA 02125, USA*
[b] *Department of Ophthalmology and Visual Sciences, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada*
[c] *Division of Epidemiology, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario M5T 3M7, Canada*

ARTICLE INFO

Article history:
Received 8 July 2020
Revised 21 March 2021

Keywords:
Crossmodal sound–shape correspondence
Sound symbolism
Developmental change
Audio–tactile (AT)
Audio–visual (AV)
multisensory development

ABSTRACT

Sound–shape crossmodal correspondence, the naturally occurring associations between abstract visual shapes and nonsense sounds, is one aspect of multisensory processing that strengthens across early childhood. Little is known regarding whether school-aged children exhibit other variants of sound–shape correspondences such as audio–tactile (AT) associations between *tactile* shapes and nonsense sounds. Based on previous research in blind individuals suggesting the role of visual experience in establishing sound–shape correspondence, we hypothesized that children would show weaker AT association than adults and that children's AT association would be enhanced with visual experience of the shapes. In Experiment 1, we showed that, when asked to match shapes explored haptically via touch to nonsense words, 6- to 8-year-olds exhibited inconsistent AT associations, whereas older children and adults exhibited the expected AT associations, despite robust audio–visual (AV) associations found across all age groups in a related study. In Experiment 2, we confirmed the role of visual experience in enhancing AT association; here, 6- to 8-year-olds could exhibit the expected AT association if first exposed to the AV condition, whereas adults showed the expected AT association irrespective of whether the AV condition was tested first or second. Our finding suggests that AT sound–shape correspondence is weak early in development relative to AV sound–shape correspondence, paralleling previous findings on the development of other types of multisensory associations. The potential role of visual experience

---

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

https://doi.org/10.1016/j.jecp.2021.105167

in the development of sound–shape correspondences in other senses is discussed.

Introduction

To function in a multisensory world, it is critical to know what sensory inputs belong together. Sensory inputs that coincide in time (temporal coincidence) or space (spatial coincidence) are more likely to be integrated (e.g., Meredith & Stein, 1983; but see Spence, 2013, for a review suggesting that spatial coincidence is not necessary for multisensory integration). In addition to temporal and spatial features, information can be linked across the senses based on semantic congruence (e.g., Doehrmann & Naumer, 2008) or abstract features (e.g., linking visual elevation and auditory pitch; Spence, 2011; Spence & Sathian, 2020). Unifying information across the senses can lead to motor and perceptual consequences such as faster responses to multisensory stimuli (e.g., Todd, 1912; Diederich & Colonius, 2004; Wang, Blohm, Huang, Boehnke, & Munoz, 2017) and the perception of multisensory illusions where information from one sense biases the processing of information in another sense (e.g., ventriloquist illusion: Vroomen & De Gelder, 2004). Understanding the typical development of multisensory processing is important given the potential benefits for perception, which supports higher cognitive functions.

Whereas the ability to use temporal synchrony and spatial coincidence to link features across the senses is present during infancy (e.g., Bahrick & Lickliter, 2000, 2004; Bremner, Mareschal, Lloyd-Fox, & Spence, 2008; Lewkowicz, 1996; Neil, Chee-Ruiter, Scheier, Lewkowicz, & Shimojo, 2006), the benefits of audio–visual (AV) integration for perception (Barutchu et al., 2019) and motor responses (Barutchu, Crewther, & Crewther, 2009; Barutchu et al., 2010; Neil et al., 2006), as well as the optimization of AV integration (Gori, Sandini, & Burr, 2012), continue to develop well into childhood. Furthermore, the development of integration between different senses does not occur at the same rate. For example, the precision of simultaneity perception between auditory white noise bursts and taps to the index finger develops more slowly than it does between auditory white noise bursts and a visual flash (Stanley, Chen, Lewis, Maurer, & Shore, 2019). Studies examining integration of sensory cues in size judgment found that whereas 8- to 10-year-old children can show partially mature visuo–tactile (VT) integration between visual and haptic cues (Gori, Del Viva, Sandini, & Burr, 2008), they do not show audio–tactile (AT) integration between auditory and haptic cues like adults (Petrini, Remark, Smith, & Nardini, 2014). Both lines of studies show that AV integration matures earlier than the integration of other sensory combinations such as AT and VT (Stanley et al., 2019). However, little is known regarding the relative development across sensory combinations of other types of multisensory perception beyond spatial/temporal coincidence and sensory cue integration.

Here we examined the development of another constraint on multisensory processing, crossmodal sound–shape correspondence. One example of crossmodal correspondence is between vision and audition, the "bouba/kiki" effect, in which an abstract visual shape (e.g., a spiky visual shape) is associated with an abstract nonsense sound (e.g., a /kiki/ or /takete/ sound) (Köhler, 1929, 1947). Variants of the bouba/kiki effect have also been found across different sensory combinations in adults (for a review, see Spence, 2011). For instance, adults associate an abstract sound like /kiki/ or /takete/ with the tactile experience of moving a robotic arm along a jagged trajectory (Fontana, 2013) and with the touch of a sharp three-dimensional (3D) object hidden from view (Fryer, Freeman, & Pring, 2014; Hamilton-Fletcher et al., 2018; Sourav et al., 2019). Furthermore, the motor system can be involved in supporting sound–shape correspondences, as shown in adults who associate heard pseudowords with seen shapes of bodily actions and associate visual shapes with heard sounds of bodily actions (Margiotoudi & Pulvermuller, 2020). Sound–shape correspondences present an interesting case study for understanding the development of associations between the senses, as elaborated below.

First, the AV bouba/kiki effect has been reported early in development (Asano et al., 2015; Maurer, Pathman, & Mondloch, 2006; Ozturk, Krehm, & Vouloumanos, 2013; Spector & Maurer, 2011), yet changes in the strength of the effect have been observed across childhood (e.g., Chow & Ciaramitaro, 2019; Pejovic & Molnar, 2017; Tzeng, Nygaard, & Namy, 2017; see also meta-analysis by Fort et al., 2018). Older children typically show larger AV effects as defined by effect size (e.g., Fort et al., 2018), sensitivity to cues that conform to the association (e.g., Tzeng et al., 2017), or the magnitude of the association compared with at-chance performance (e.g., Chow & Ciaramitaro, 2019). Compared with the wealth of studies on the development of the AV bouba/kiki effect, much less is known about the development of the AT bouba/kiki effect. It has yet to be established whether children associate nonsense words and tactile shapes as adults do using similar AV testing paradigms. A relevant study reported comparable association strengths for a different crossmodal correspondence, changing auditory pitch and visual movement, in 4- and 5-year-old preschoolers (Nava, Grassi, & Turati, 2016). Based solely on this study, however, it is hard to generalize these findings from pitch–motion correspondence using passive tactile exploration to the crossmodal correspondence we are examining, sound–shape correspondence using active tactile exploration. In other studies on multisensory processing, AT simultaneity judgments were found to be adult-like in 11-year-olds (e.g., Stanley et al., 2019), which is delayed compared with adult-like AV simultaneity judgment in 9-year-olds (Chen, Shore, Lewis, & Maurer, 2016). Given this previous work, we hypothesized that children younger than 11 years would show weaker AT associations than AV associations.

Second, emerging literature points to the effect of early visual experience on the development of AT sound–shape correspondence. Congenitally blind and early-blind adults do not exhibit the expected AT bouba/kiki associations (Fryer et al., 2014; Hamilton-Fletcher et al., 2018; Sourav et al., 2019). Sourav et al. (2019) showed a protracted sensitive period for sound–shape correspondence; individuals with a history of developmental cataracts before 12 years of age did not show the expected AT sound–shape correspondences after cataract removal surgery, whereas late permanently blind individuals with blindness onset after 12 years of age showed comparable strength of sound–shape correspondence. If the effect of visual experience from studies in the blind can be extended to the developmental literature, we hypothesized that (1) older participants would show more robust AT sound–shape correspondence than younger participants because they have had more years of exploring shapes via touch as well as more years of visual experience and (2) younger participants lacking direct visual experience of the abstract shapes would show reduced AT sound–shape correspondences compared with younger participants having direct visual experience of the abstract shapes.

The current study

The current study aimed to understand the development of sound–shape correspondences across AV and AT shapes in 6- to 35-year-old typically developing individuals. In two experiments we quantified the strength of AT and AV associations between abstract shapes and nonsense words in children as young as 6 years and in adults. Previous research suggests that haptic object recognition improves from 3 to 12 years of age (Dunn et al., 2013; see review by Taylor et al., 2016). Nonetheless, our child participants at 6 years of age or older should be able to explore an object shape via touch, as has been shown in 5-year-olds (Kalagher & Jones, 2011). In Experiment 1, we used a between-participants design and tested whether children exhibited the expected AT association and compared the strength of AT association with that of a previously reported AV association in a different group of observers (Chow & Ciaramitaro, 2019). In Experiment 2, we used a within-participant design to compare the strength of AT and AV association in the same observers (6- to 8-year-olds and adults) and examined the effect of visual experience by having half of the observers complete the AV condition first and the other half complete the AT condition first. We expected to find that (1) AT associations would develop more slowly than AV associations, in accord with previous literature on the development of spatial/temporal integration of multisensory stimuli, (2) the youngest age group (6- to 8-year-olds) would show weaker association strength regardless of modality compared with the older children and young adults, in accord with previous developmental literature on AV sound–shape correspondences (e.g., Chow & Ciaramitaro, 2019), and (3) AT and AV association strength would be more comparable in 6- to 8-year-olds when the abstract shapes were seen prior to AT.

Experiment 1: Development of AV and AT correspondences across participants

Method

Participants

A total of 21 adults (47.6% female; 18–35 years of age) and 63 children (55.6% female; 6–17 years of age) were recruited to participate in the AT condition at the Living Laboratory @ the Museum of Science, Boston, a large city in the northeastern United States (see Table 1 for detailed demographics). We also included data from 21 adults (38.1% female; 18–35 years of age) and 47 children (53.2% female; 6–17 years of age) recruited in a complementary AV condition (for details, see Table 1) (Chow & Ciaramitaro, 2019). We set a minimal sample size of 20 participants for the age groups 6 to 8, 9 to 11, and 18 to 35 years based on previous studies using similar designs (e.g., Chow & Ciaramitaro, 2019, n = 14–21 per age group; Maurer et al., 2006, n = 20 per age group). A smaller number of participants was recruited for the adolescent age group (12–17 years, n = 12) based on previous findings that the association plateaued before adolescence (Chow & Ciaramitaro, 2019). Adult participants gave written consent. Legal guardians of children under 18 years of age gave written permission, and children under 18 years of age gave verbal assent and written consent when possible. All experimental protocols were approved by the institutional review board of the University of Massachusetts Boston and of the Living Laboratory @ the Museum of Science, Boston, and complied with the Declaration of Helsinki.

Stimuli and apparatus

For both the AV and AT conditions, visual and auditory stimulus presentation was controlled by MATLAB (R2014b or 2009b) and Psychtoolbox Version 3.0.12 (Brainard, 1997; Pelli, 1997) run on a MacBook Pro (15 inches). Four nonsense sounds (/baba/, /gaga/, /kiki/, and /titi/) were recorded using the text-to-speech function of Mac OS X and edited with Audacity (Audacity Team, 2014) to standardize sound duration and amplitude across sound files. Auditory stimuli were presented via noise-canceling headphones (3 M Peltor, HTB79A). Pairs of visual stimuli were presented on a computer display for the AV condition. Pairs of tactile stimuli were handed to the participant in the AT condition hidden from view. Responses were directly made either by the participant in the AV condition or by the experimenter in the AT condition via button press on a Cedrus response pad (RB-844) and were recorded by MATLAB.

For the AV condition, eight abstract visual shapes were created in Adobe Illustrator varying in the number of protrusions (three or seven), size of protrusions (small or large), and contour spikiness (round or spiky). These features were similar to features investigated using radial frequency patterns (frequency, amplitude, and spikiness) to study AV bouba/kiki associations (Chen, Huang, Woods, & Spence, 2016, 2019). We presented four repeats of each AV stimulus combination for a total of 64 trials (4 pairs of shapes × 4 nonsense sounds × 4 repeats). However, only a subset of data (seven-protrusion

Table 1
Participant demographics in Experiment 1 (AT condition; N = 84).

			Within age group			
Characteristic	*n*	% of *N*	% White	% English as (one of) first language(s)	% Spanish as (one of) first language(s)	% Bilingual
Female	45	53.6				
Age group						
6–8 years	21	25.0	66.7	100.0	9.5	28.6
9–11 years	30	35.7	73.3	86.7	6.7	23.3
12–17 years	12	14.3	58.3	100.0	8.3	25.0
18–35 years	21	25.0	76.2	85.7	9.5	19.0

shape comparisons; 32 trials) was analyzed in the current study to allow comparison across the AV and AT conditions. For further details about the AV condition, see Chow and Ciaramitaro (2019).

For the AT condition, four tactile shapes were used, each with a maximum distance from one protrusion tip to the opposite protrusion tip of 10 cm. Shapes were made out of molding foam or clay to resemble the seven-protrusion visual shapes used in the AV condition (Chow & Ciaramitaro, 2019). Stimulus thickness was consistent (1 cm) across all shapes, keeping this one shape dimension constant and creating two-dimensional (2D) tactile objects varying in shape contour. Stimulus texture and material composition were also held constant. Two shapes were spiky with seven *pointed* ends extending outward (one with thin protrusions and the other with thick protrusions). Two shapes were complementary rounded versions of the spiky shapes, with seven *rounded* ends extending outward (one with thin protrusions and the other with thick protrusions). See Fig. 1 for sample foam stimuli and display apparatus. Shapes with similar protrusion size, thick or thin, were always presented as a pair to participants. Shapes were presented hidden from view inside a cardboard box (44.7 × 29.4 × 24.4 cm). The front of the box, the participant's view, contained two openings (14.7 × 9.5 cm) covered by cloth curtains (see Fig. 1B). The back of the box, the experimenter's view, contained a large opening to allow the experimenter to place the shapes in and switch them out between trials (see Fig. 1C). We presented four repeats of each AT stimulus combination (2 pairs of shapes × 4 nonsense sounds) for a total of 32 trials.

Procedure

In the AV condition, participants viewed two visual shapes and heard one nonsense word (/baba/, /gaga/, /kiki/, or /titi/; 700 ms each) and were instructed to choose the shape that matched the sound by pressing the button on the corresponding side (for details, see Chow & Ciaramitaro, 2019). The procedure for the AT condition was similar to that for the AV condition except that shapes were felt and hidden from view. Participants sat in front of the cardboard box and, upon hearing an auditory cue, placed each hand into a given side of the box and felt the shapes the experimenter had placed inside, one shape for each hand on each side. While the participant felt the shapes, a nonsense word was

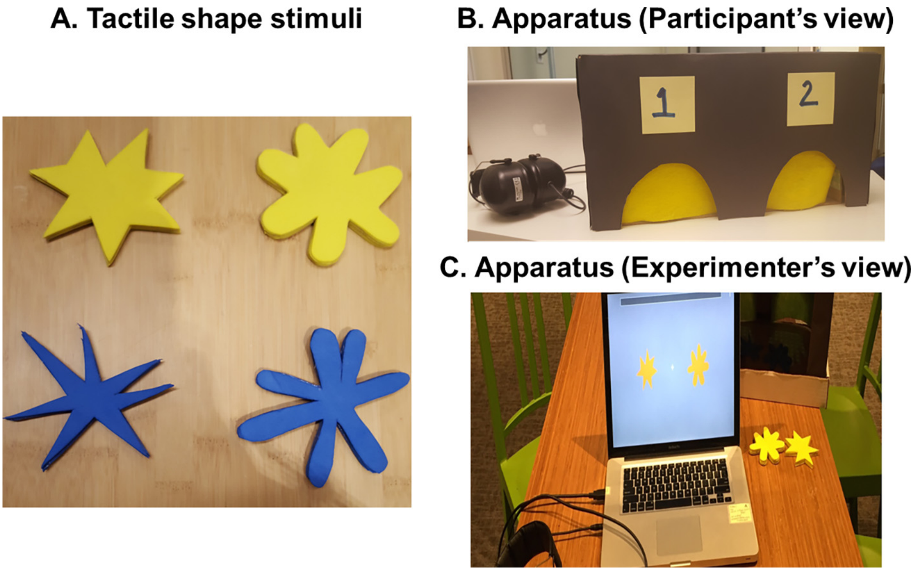

Fig. 1. (A) Tactile stimuli made out of foam. The top round/spiky shape pair has thicker protrusions, whereas the bottom round/spiky shape pair has thinner protrusions. Shape pairs were presented such that protrusion size was always matched and the only difference between shape pairs was contour spikiness (round vs. spiky, as shown on the right and left, respectively). (B,C) Apparatus used in the study from the participant's view (B), with the box hiding the shapes from view, and from the experimenter's view (C), with the computer screen showing the shape pair to present.

presented via headphones for 700 ms (Fig. 2). Participants indicated which shape they judged to match the sound by verbally indicating the location of the shape (left or right) or by lifting the chosen shape inside the box such that only the experimenter could see the choice and record the response via button press. After the response was entered, participants heard a second auditory cue, which indicated that the trial was over and hands should be removed from the box to proceed to the next trial. If participants did not respond within 30 s, the trial was aborted and repeated. To keep participants motivated, a status bar incremented to indicate trials completed at the half-way and three-quarter-way points of the experiment.

To ensure that performance differences did not arise from a lack of understanding or from sound levels not being adequate across participants, practice trials were completed before the experiment began. In practice trials, participants either matched cartoon animal images (cat vs. dog) to an animal sound (/mjaʊ/ or /wʊf/) in the AV condition or matched toys (stuffed elephant vs. toy car) hidden from view to a sound (elephant making noise or car honking) responding as described above. Participants needed to complete 4 consecutive practice trials correctly. If participants failed to reach this criterion in 20 practice trials, their data were discarded from further analysis.

Measures

Binary choice response (0 or 1). Each trial's response was coded as 0 if the participant chose a spiky shape or 1 if the participant chose a round shape.

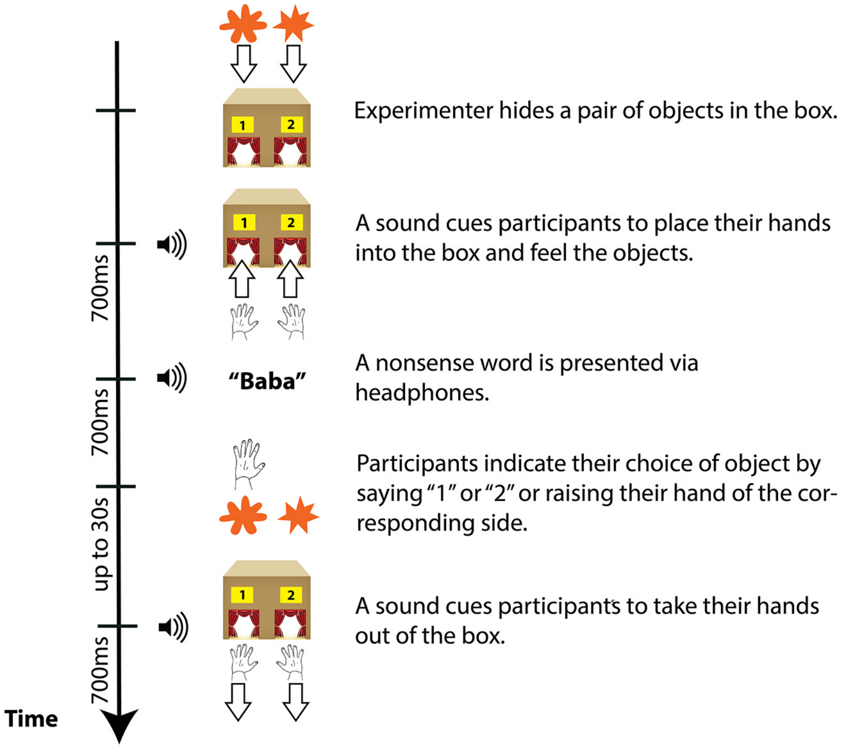

Fig. 2. Procedure for a sample trial. A sound cued participants to place their hands in the box, and then participants heard a nonsense word—/baba/, /gaga/, /titi/ or /kiki (700 ms)—while feeling two opposing tactile shape pairs and had up to 30 s to make a choice. After participants indicated their choice verbally by saying “1″ or “2” or by raising their hand on the corresponding side where they felt the matching shape, a sound cued them to take their hands from the box.

Choice bias separated by sound type (range = − 0.5 to 0.5). To detect the presence or absence of associations across trials, a choice bias was computed with the proportion of trials where the choice response was 1 (round shape/object chosen) for a given sound (j) by a given observer (i) for AV and AT associations, respectively. If participants made their choice randomly, this proportion would be 0.5. Thus, by subtracting 0.5 from this proportion (Eqs. (1) and (2)), we were able to get a round bias measure with 0 reflecting no bias [choice bias$_{(i,j)}$]. A positive choice bias indicates a round shape is preferred, whereas a negative choice bias indicates a spiky shape is preferred. We averaged choice bias across a given vowel; for example, choice bias for /baba/ and /gaga/ sounds were collapsed together to indicate an overall choice bias for /a/ sounds (see Supplementary Fig. 1 in online supplementary material):

$$\text{AT choice bias}_{(i,j)} = P(\text{round object chosen})_{(i,j)} - 0.5 \tag{1}$$

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

We computed the above measures based on a subset of the results from Experiment 2 by Chow and Ciaramitaro (2019), including only responses for abstract visual shapes analogous to the abstract tactile shapes used in the current study (shapes with seven large or small protrusions). Thus, any differences between AV and AT correspondences should not be due to differences in shape stimuli.

Data analysis

All analyses described below were performed in R (R Core Team, 2017).

Presence/Absence of association. We compared the choice bias against zero to establish the presence of an association in each age group and for each modality tested using the *t.test* and *wilcox.test* functions. The p values were adjusted for multiple comparisons using the *p.adjust* function to control for false discovery rate (Benjamini & Hochberg, 1995). The effect size was denoted by Hedges' g (Cohen's d corrected for small sample size) and estimated using the R package "effsize" (Torchiano, 2017). Results were similar across t tests and Wilcox tests, so t-test results were reported for easier comparison with previous literature.

Magnitude of association. We also modeled changes in the magnitude of crossmodal associations across development and modalities with a generalized linear mixed-effect model using participants' binary choice response using the R package "lme4" (Bates, Maechler, Bolker & Walker, 2015) *glmer* function. To compare the magnitude of crossmodal correspondence *across development and modalities,* we performed mixed-effect logistic regression modeling to predict whether a round shape is chosen over a spiky shape, with the following fixed-effect factors: sound vowel category (/a/ or /i/), age group (6- to 8-year-olds or 18- to 35-year-olds), and tested modality (AT or AV) and their interaction terms, as well as participant as a random effect (intercept), with a binomial distribution (Laplace approximation). To simplify comparisons, we included data from only younger children (6- to 8-year-olds) and adults (18- to 35-year-olds) and dichotomized the variable into young children or adults. We expected to see a significant interaction between sound vowel category and age group if the magnitude of crossmodal correspondence changes across development. Likewise, we expected to see a significant effect of modality and an interaction between modality and age group if the magnitude of the crossmodal correspondence changes across modalities. The final model included 2415 raw choice responses collected from 68 participants.

Results

Our final sample consisted of 84 usable participants for the AT condition (the current study) and 68 usable participants for the AV condition (data from Chow & Ciaramitaro, 2019). To be included, data needed to meet the same criterion used in our previous study (Chow & Ciaramitaro, 2019). Data were discarded from an additional 11 participants who failed to reach criterion during the practice phase after the maximum of 20 possible practice trials (n = 4), failed to complete at least 8 trials (n = 4), or did not fill in the demographic questionnaires properly (n = 3). Participants included for analysis completed 4 (the minimum number) to 8 practice trials before reaching practice criterion. Out of a maximum of 32 possible trials, participants completed an average of 30.3 trials. The average number of trials broken down by age is included in Fig. 3.

Presence or absence of the AV and AT associations across development

Fig. 3 plots the average AT choice bias for different age groups (Panel A) and the complementary AV choice bias in different participants for the same age groups (Panel B). A positive mean indicates stronger preferences for correspondences with more rounded objects, whereas a negative mean indicates stronger preferences for correspondences with more spiky objects. We found an insignificant AT choice bias in 6- to 8-year-olds for both /a/-related sounds, $t(20) = 1.724$, $p = .114$, $g = 0.369$, and /i/-related sounds, $t(20) = -1.208$, $p = .241$, $g = -0.259$. On the contrary, the AV choice bias in another group of 6- to 8-year-olds was significantly above zero for /a/-related sounds, $t(14) = 3.536$, $p = .004$, $g = 0.888$, and close to being significantly negative for /i/-related sounds, $t(14) = -2.141$, $p = .050$, $g = -0.538$. Both the AT and AV choice biases for /a/- and /i/-related sounds were significantly different from zero in the expected directions in all other age groups ($ts \geq 2.301$, $ps \leq .038$, $gs \geq 0.415$).

Magnitude of the AV and AT associations across development

We performed mixed-effect binary/logistic regression modeling to predict participants' odds of choosing a round shape when a given sound was presented, when a given modality was tested (AT or AV), and when participants were from a given age group (6–8 years or 18–35 years) (Akaike information criterion [AIC] = 2811.9, Bayesian information criterion [BIC] = 2864.0, log likelihood = −1397.0). Table 2 lists the odds ratio coefficient of each predictor in the alternative model. Two interesting effects of tested modality emerge. First, the model reports a significant interaction between sound vowel category and tested modality; the odds of choosing a round shape/object when an /a/ sound was presented increased by a factor of 3.60 ($p < .001$) when participants were tested by sight versus touch, holding other factors constant. Second, the model reports a significant three-way interaction among age group, sound category, and tested modality; in adults, the odds of choosing a round shape when an /a/ sound was presented decreased by a factor of 0.36 ($p = .009$) when participants were tested by sight versus touch, holding other factors constant. This suggests that the discrepancy in association strength across modalities is reduced in adults compared with 6- to 8-year-olds.

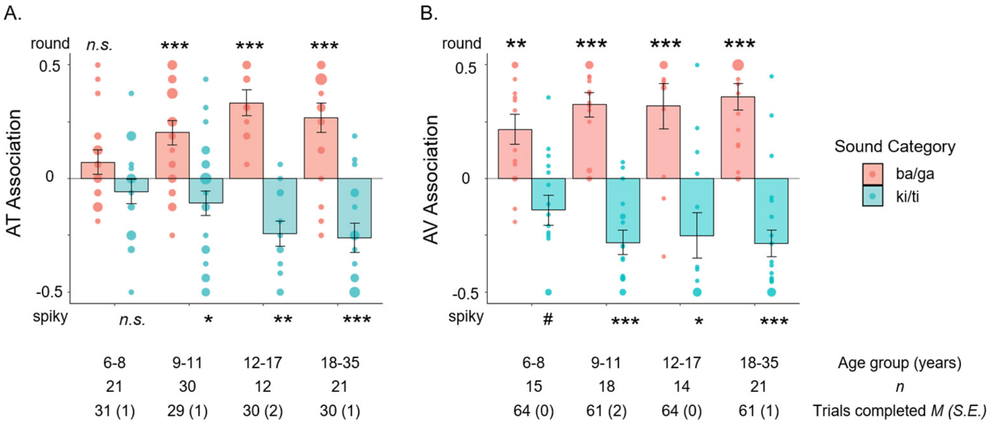

Fig. 3. Individual audio–tactile (AT) choice bias (dots with size scaled to account for multiple participants at the same value) and mean AT choice bias (bar plot ± standard errors across participants) results from Experiment 1 show that younger children (6- to 8-year-olds) did not associate nonsense words and abstract tactile shapes (A), unlike another group of participants of the same age showing a significant association between nonsense words and the same abstract shapes when presented visually (B), a subset of data from a different group of participants tested in Chow and Ciaramitaro (2019). Both AT and audio–visual (AV) associations were above chance for older children and young adults. *n.s.*, nonsignificant; #$p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$.

Table 2
Fixed effects of the final mixed-effect logistic regression model predicting binary response using age group, sound category, and tested modality as predictors based on data from Experiment 1.

Fixed effects	Odds ratio	*SE*	*z* Value	*p* Value	95% Confidence interval	
					Lower	Upper
(Intercept)	0.859	0.165	−0.921	.357	0.622	1.186
Age group						
1: 18–35 years	0.367	0.243	−4.128	<.001	0.228	0.591
0: 6–8 years						
Sound category						
1: /baba/ or /gaga/	1.572	0.163	2.782	.005	1.143	2.163
0: /kiki/ or /titi/						
Tested modality						
1: sight (AV)	0.569	0.229	−2.460	.014	0.363	0.892
0: touch (AT)						
Age Group × Sound Category	7.991	0.254	8.183	<.001	4.857	13.145
Age Group × Tested Modality	1.586	0.326	1.415	.157	0.838	3.002
Sound Category × Tested Modality	3.602	0.266	4.812	<.001	2.137	6.071
Age Group × Sound Category × Tested Modality	0.360	0.387	−2.644	.008	0.169	0.768

Note. Number of observations: 2415 grouped by participant ID (*n* = 68). Random effects: participant (intercept) variance = 0.3427, *SD* = 0.5854. AT, audio–tactile (data from the current study); AV, audio–visual (a subset of data from Chow & Ciaramitaro, 2019).

Discussion

We found important differences between this new study, examining AT associations, and our previous study, examining AV associations (data from Chow & Ciaramitaro, 2019). Our youngest participants, 6- to 8-year-olds, showed no significant sound–shape correspondence when tested by touch (AT), even though children of the same age showed significant sound–shape correspondence when tested by vision (AV) in a different study (Chow & Ciaramitaro, 2019). On the contrary, adults showed comparable and significant AT and AV associations, both of which were stronger than associations found in 6- to 8-year-olds.

We consider two primary factors that could limit AT associations but not AV associations. First, the strength of AT associations might be limited by touch-specific factors not shared with AV associations. In addition to failing to make systematic matches in a tactile match-to-sample task (Kalagher & Jones, 2011), younger children are less sophisticated in haptic exploration, using more exploratory procedures but not necessarily the most efficient ones (Withagen, Kappers, Vervloed, Knoors, & Verhoeven, 2013). As a result, haptic judgment is generally less accurate (Alexander, Johnson, & Schreiber, 2002) and requires longer exploration times (Withagen et al., 2013). Furthermore, children's haptic performance in estimating shape is particularly impoverished compared with their ability to estimate other tactile dimensions such as weight, texture, and volume (Withagen et al., 2012). Even when using the same exploratory procedures, younger children tend to prioritize information other than shape, such as texture, when using tactile exploration (Lederman & Klatzky, 1987; Schwarzer, Küfer, & Wilkening, 1999). Although some research suggests that 6- to 8-year-olds could exhibit effective haptic exploration strategies, it is possible that with the unfamiliar abstract shapes used in the current experiment, children required further instruction to understand the goal of exploration and adopt the most effective haptic exploration strategy.

If weaker AT associations in the current experiment arise from ineffective haptic exploration strategies in 6- to 8-year-olds, explicit instruction of the most effective haptic exploration strategy directed at the contour of the shapes should lead to stronger AT associations, especially in this young age group. Thus, in Experiment 2, we provided explicit instructions and focused on comparing the performance of 6- to 8-year-olds and 18- to 35-year-olds. Prior to the experiment, participants watched an instructional video highlighting contour following (i.e., smooth and nonrepetitive movement along the object contour; Lederman & Klatzky, 1987). We expected that a more explicit instruction of haptic exploration strategy would allow a fairer comparison of association strength across age groups.

Second, the strength of AT associations might be limited by visual factors. To understand this relationship, in Experiment 2 the same participants completed both the AT and AV blocks with testing order counterbalanced. This within-participant design allowed us to consider association strength in the same participant to take into account individual differences and to compare association strength based on whether the AV block was tested first, allowing a direct investigation of the role of visual experience.

Experiment 2: Development of AV and AT correspondences within participants, the role of explicit instruction for haptic exploration, and the role of visual experience

Method

Participants

A total of 30 adults (76.7% female; 18–35 years of age) and 30 children (66.7% female; 6–8 years of age) were recruited and consented to participate in this experiment at the Living Laboratory @ the Museum of Science, Boston (see Table 3 for detailed demographics).

Apparatus and stimuli

The apparatus and stimuli were similar to those used in Experiment 1 with the following exceptions. First, the tactile shapes were 3D printed in the current experiment and were sturdier and resembled the visual shapes perfectly. The 3D design of the shapes was created by adding depth (1 cm) to the original visual abstract shapes using Tinkercad (https://www.tinkercad.com). Second, the box in which the tactile shapes were presented was replaced by a taller box (height = 45 cm) to allow better lighting for video-recording of participants' hand movements while blocking participants' view of stimuli. Third, Velcro was added to shapes to hold them in place better and to allow for easier simultaneous haptic exploration with both hands, one shape in each hand. Fourth, a video camera (Logitech C920 HD Pro Webcam) was used to record participants' hand movements during the AT experiment to allow offline coding of haptic exploration strategies.

Procedure

Participants completed two blocks; one block tested AT association, where they chose which of two felt but not seen objects best matched a sound, and one block tested AV association, where they chose which of two seen visual shapes best matched a sound. Block order was counterbalanced across participants such that half of the participants in each age group did AT first and half did AV first. Two repeats of each AT stimulus combination (2 pairs of shapes × 4 nonsense sounds) were presented for a total of 16 trials per block.

The procedure for AV blocks was similar to that described in a previous article (Chow & Ciaramitaro, 2019), and the procedure for AT blocks was similar to that in Experiment 1 except that, prior to starting the block, participants were shown an instructional video (see Supplementary Fig. 2 for details) and asked to demonstrate the haptic exploration strategy, contour following, with practice objects (a cutout shape of a cartoon elephant and a car). After training, the experimenter explained the task in a similar way as in Experiment 1 and started video-recording participants' hand movements.

Table 3
Participant demographics in Experiment 2 (N = 60).

			Within age group			
Characteristic	*n*	% of *N*	% White	% English as (one of) first language(s)	% Spanish as (one of) first language(s)	% Bilingual
Female	43	71.7				
Age group						
6–8 years	30	50.0	80.0	90.0	3.3	33.3
18–35 years	30	50.0	56.7	70.0	20.0	33.3

Measures

Binary choice response (0 or 1). This was the same as in Experiment 1.

Choice bias separated by sound type (range = – 0.5 to 0.5). This was the same as in Experiment 1.

Visual enhancement index (range = – 2 to 2). To quantify how AV association might be stronger than AT association, for each participant we computed a visual enhancement index by subtracting the overall choice bias of AT from that of AV (Eq. (3)). The overall choice bias for each modality was computed by subtracting choice bias of /i/-related sounds from that of /a/-related sounds. If a participant exhibited the expected association across the two types of sounds (choice bias for /a/-related sounds is 0.5 and that for /i/-related sounds is −0.5), the overall choice bias would be 1; if the participant exhibited the opposite association (choice bias for /a/-related sounds is −0.5 and that for /i/-related sounds is 0.5), the overall choice bias would be −1 (for further analyses, see Supplementary Fig. 3). A zero score for the visual enhancement index indicates no difference in overall association strength between AT and AV. A positive score indicates stronger AV association than AT association in the expected direction, and vice versa for a negative score:

$$\begin{aligned}\text{Visual enhancement index}_{(j)} = {} & \left(\text{AV choice bias}_{(/\mathrm{a}/,j)} - \text{AV choice bias}_{(/\mathrm{i}/,j)}\right) \\ & - \left(\text{AV choice bias}_{(/\mathrm{a}/,j)} - \text{AV choice bias}_{(/\mathrm{i}/,j)}\right) \qquad (3)\end{aligned}$$

Data analysis

Direction and magnitude of AV and AT associations. This was the same as in Experiment 1.

Visual enhancement effects. To examine the influence of AV association on AT association, we compared the visual enhancement index for each age group (6- to 8-year-olds or 18- to 35-year-olds) and testing order (AT tested before or after AV) against zero using the R package "stats" *t.test* and *wilcox.test* functions and reported Hedges' *g* (Cohen's *d* corrected for small sample size) estimated by the R package "effsize" (Torchiano, 2017) *cohen.d* function. The *p* values were adjusted for multiple comparisons using the *p.adjust* function, controlling for false discovery rate (Benjamini & Hochberg, 1995).

Results

Our final sample consisted of 60 usable participants. Data were discarded from an additional 16 participants who failed to complete all trials in both the AT and AV blocks (n = 12) or whose hand movements were not observable because the video camera was not placed properly (n = 4). All participants completed 4 (the minimum number) to 8 practice trials before reaching practice criterion; no participants were discarded because of failure to reach practice criterion within 20 practice trials. All participants included in the final sample, regardless of age, completed all 32 possible test trials.

Presence or absence of AV and AT associations

Fig. 4 plots the average AT choice bias for young children and adults (Panel A, C) and the complementary AV choice bias in the same participants (Panel B, D), separated by whether AT was tested before AV (top panel: A, B) or after AV (bottom panel: C, D). Similar to the results from Experiment 1 and previous reports, adult participants showed significant AT and AV choice biases in the expected direction for both sound types. This effect is not modulated by testing order, $ts(14) \geq 5.277$, $ps \leq .001$, $gs \geq 1.326$.

We found a significant AT choice bias in 6- to 8-year-olds for both /a/-related sounds, $t(14) = 2.667$, $p = .032$, $g = 0.670$, and /i/-related sounds, $t(14) = -2.428$, $p = .038$, $g = -0.610$, but only when the AT block was tested *after* the AV block. When the AT block was tested *before* the AV block in a different group of 6- to 8-year-olds, we found no significant AT choice bias for either /a/-related sounds, $t(14) = 0.627$, $p = .541$, $g = 0.158$, or /i/-related sounds, $t(14) = 0.705$, $p = .525$, $g = 0.178$. This suggests that explicitly instructing children to explore the shape contour alone does not allow for the exhibition of AT association.

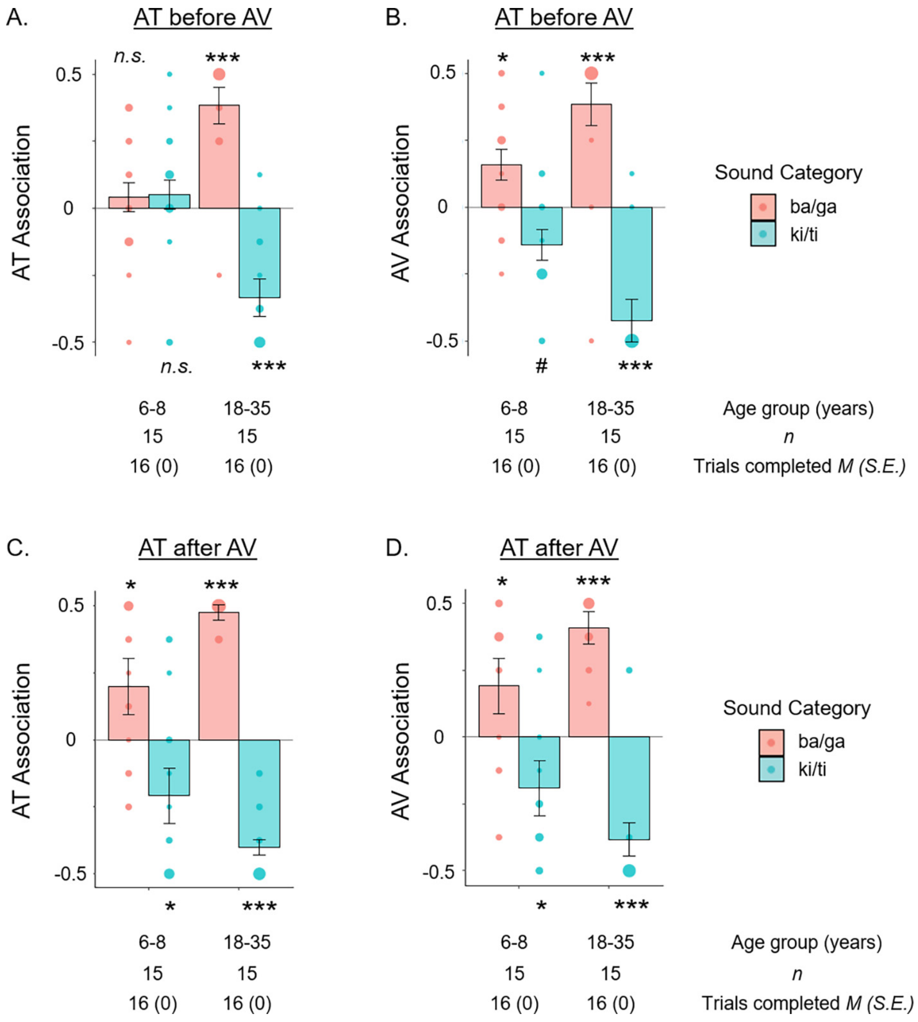

Fig. 4. Individual audio–tactile (AT) choice bias (dots with size scaled to account for multiple participants at the same value) and mean AT choice bias (bar plots ± standard errors across participants) results from Experiment 2 show that younger children (6- to 8-year-olds) demonstrated the expected AT association only when tested after audio–visual (AV) (C), as opposed to showing the expected AV association regardless of the testing order (B,D). By contrast, the young adult age group showed consistent associations regardless of whether the abstract shapes were presented via vision or touch or whether AT was tested before or after AV. *n.s.*, nonsignificant; #$p < .10$; *$p < .05$; ***$p < .001$.

Unlike the results for AT choice bias, AV choice bias in 6- to 8-year-olds was significantly different from zero for /a/-related sounds, $ts(14) \geq 2.400$, $ps \leq .038$, $gs \geq 0.603$, and was marginally different from zero for /i/-related sounds, $ts(14) \geq -2.125$, $ps \leq .059$, $gs \geq 0.534$, regardless of the testing order. The relatively more robust AV association for /a/-related sounds and rounded shapes shown here is consistent with what was reported in Chow and Ciaramitaro (2019) and Fort et al. (2018).

Magnitude of AV and AT associations

We performed mixed-effect binary/logistic regression modeling to predict participants' odds of choosing a round object (AT) or round shape (AV) when a given sound was presented, including 1920 raw binary choice responses collected from 60 participants. Similar to Experiment 1, we included both AT and AV data in the same logistic model and reported the final model, which included tested modality as a predictor, along with age group and sound vowel category (AIC = 1869.7, BIC = 1919.7, log likelihood = − 925.83). Unlike the analysis for Experiment 1, we also added testing order (AT tested before or after AV) as a predictor. We expected that if the association was consistent across tested modality, tested modality would not be a significant predictor. In addition, if the association was bidirectional, the testing order should not be a significant predictor. Table 4 lists the odds ratio of each predictor in the final model. Similar to the results in Experiment 1, the model reports a significant interaction between sound vowel category and tested modality; the odds of choosing a round shape when an /a/ sound was presented increased by a factor of 1.90 (p = .020) when participants were tested by sight versus touch, holding other factors constant. In contrast to Experiment 1, in Experiment 2 we did not find a three-way interaction across age group, sound category, and tested modality (p = .187); the effect of tested modality on association strength was consistent across tested age groups. However, we noted a stronger within-participant correlation between AT and AV association magnitude in both age groups when AT was tested after AV (see Supplementary Fig. 3).

Enhanced AV association

Fig. 5 plots the average visual enhancement index across age groups and testing order (individual data are indicated by dots). Mean visual enhancement index was significantly different from zero for children when AT was tested before AV (M = 0.308), $t(14)$ = 4.054, p = .005, g = 1.019, but not when AT was tested after AV (M = − 0.025), $t(14)$ = − 0.262, p = .797, g = − 0.066. Mean visual enhancement was not significantly different from zero in adults in either testing order ($ts \leq 1.435$, $ps \geq .387$, $gs \leq 0.360$). These results confirm that the enhanced AV association observed in previous analyses is specific to children, not adults, and specific to the testing order, seen only when AT was tested before AV. Our results suggest that some aspect of AV testing enhances subsequent AT associations during AT testing.

General discussion

The goal of the current study was to understand the development of AT associations between abstract shapes and nonsense words, and to consider how they compare with AV associations. This work allows us to address a broader question regarding the development of AT processing relative to AV processing. In Experiment 1, we measured the strength of AT associations across a wide age

Table 4

Fixed-effect coefficients of the final mixed effect logistic regression model predicting binary response using age group, sound category, and tested modality as predictors based on data from Experiment 2.

Fixed effects	Odds ratio	*SE*	*z* Value	*p* Value	95% Confidence interval	
					Lower	Upper
(Intercept)	0.716	0.166	−2.013	.044	0.518	0.991
Age group						
1: 18–35 years	0.192	0.276	−5.997	.000	0.112	0.329
0: 6–8 years						
Sound category						
1: /baba/ or /gaga/	2.356	0.192	4.469	.000	1.618	3.432
0: /kiki/ or /titi/						
Tested modality						
1: sight (AV)	0.674	0.195	−2.028	.043	0.460	0.987
0: touch (AT)						
Age Group × Sound Category	46.219	0.382	10.040	<.001	21.868	97.683
Age Group × Tested Modality	1.009	0.353	0.024	.980	0.505	2.016
Sound Category × Tested Modality	1.904	0.276	2.331	.020	1.108	3.273
Age Group × Sound Category × Tested Modality	0.501	0.523	−1.320	.187	0.180	1.398

Note. Number of observations: 1920 grouped by participant ID (n = 60). Random effects: participant (intercept) variance = 0.2812, SD = 0.5302.

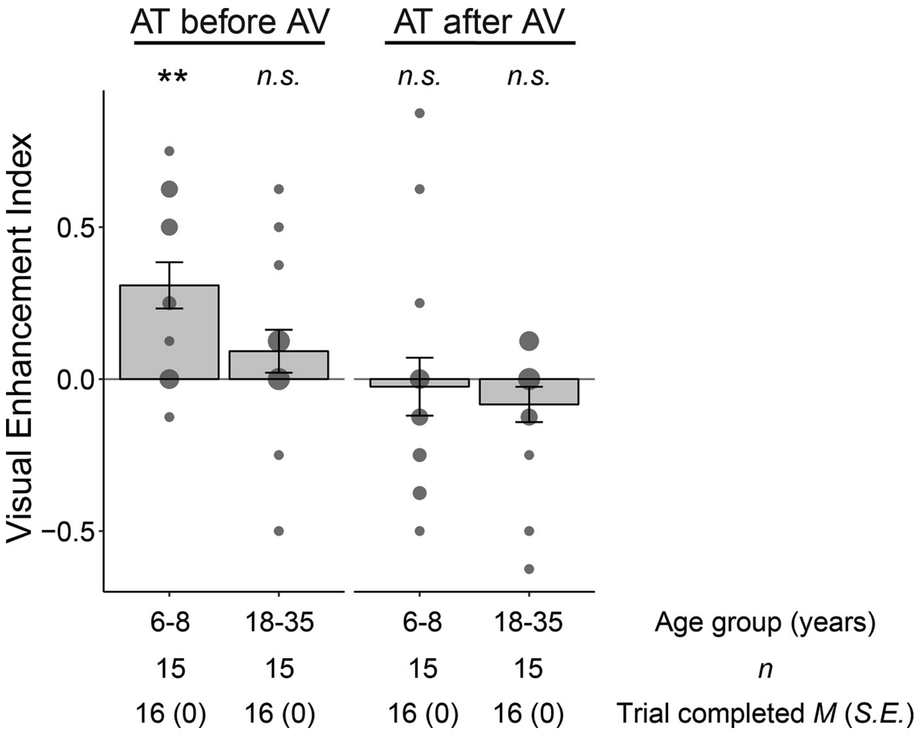

Fig. 5. Individual visual enhancement index (dots with size scaled to account for multiple participants at the same value) and mean visual enhancement index (bar plots ± standard errors across participants) between age groups and testing order in Experiment 2. Audio–visual (AV) association was enhanced relative to audio–tactile (AT) association only in children and only when AT was tested before AV. *n.s.*, nonsignificant; ***p* < .01.

range of participants from 6 to 35 years. We found a protracted development for AT associations compared with AV associations. AT associations were weak and at chance in 6- to 8-year-olds, whereas children of the same age run in a complementary paradigm showed the expected AV associations in a previous study (Chow & Ciaramitaro, 2019). Similar results were found in Experiment 2 when AT was tested before AV, even after children were taught one of the most effective haptic exploration strategies for shape judgment—contour following. This suggests that explicit instruction on haptic exploration strategy alone is not sufficient to support AT associations in children of this age. Only when AT was tested after AV, were AV and AT association strengths found to be similar within the same 6- to 8-year-olds. In contrast, adults showed similar AV and AT association strengths regardless of test order. Furthermore, in both experiments, adults exhibited stronger associations than younger children. The effect of testing order and changes across development highlight the potential role of visual experience in supporting AT sound–shape correspondence.

General development of sound–shape correspondences

Regardless of the tested modality, we found that sound–shape correspondences were weaker overall, with more diverse responses (greater individual differences) as well as stronger order effects in 6- to 8-year-olds compared with adults. This finding accords with previous literature suggesting that adults have more consistent AV mappings (i.e., stronger associations) compared with young children (Chow & Ciaramitaro, 2019; Fernández-Prieto, Navarra & Pons, 2015; Maurer et al., 2006; Ozturk et al., 2013; Tzeng et al., 2017). Furthermore, this suggests a role for experience in strengthening sound–shape correspondences into early adulthood for both AV and AT associations. We also found that, regardless of tested modality, associations between round shapes and /a/-related sounds emerge earlier than associations between spiky shapes and /i/-related sounds (AV: 6- to 8-year-olds; AT: 9- to 11-year-olds, Experiment 1). Such asymmetries in the early emergence of sound–shape correspondences were also evident in a meta-analysis examining AV sound–shape correspondence (Fort et al., 2018), which found that associations with /a/ sounds, but not /i/ sounds, appeared earlier—present by the first year of life. Our findings confirm and extend this previous work, revealing the similarities in developmental patterns for AV and AT associations.

Protracted development for AT associations relative to AV associations

We found that AT sound–shape associations develop at a more protracted rate compared with AV sound–shape associations if young participants have no prior experience with seeing the unique abstract shapes before tactile testing (Experiment 1 and Experiment 2–ATAV condition [AT before AV]). Such protracted AT development is not due to not knowing the best tactile exploration strategy given that we found that AT association strength remained at chance levels even when children were given explicit instructions and taught how to follow shape contour and were able to successfully exhibit contour following on practice trials. We acknowledge that it is possible that our younger participants still failed to explore abstract shapes effectively, which would require trial-by-trial video coding of haptic exploration strategies to validate. Nevertheless, if age and haptic shape exploration were indeed the limiting factors of AT association, we would have expected younger participants to fail to exhibit AT association even after seeing the shapes (Experiment 2–AVAT condition [AV before AT]), which was not the case. Below we elaborate on a possible interpretation of our results.

One factor contributing to our results could be inherent differences in processing styles between vision and touch. For example, young children might require a longer time per trial to extract shape contour via touch (a sequential process that requires the integration of local elements) versus via sight (a more immediate process at the global level) compared with adults. Previous research has found more comparable recognition performance across vision and touch when the exploration strategy and spatial resolution are matched by blurring visual information and revealing visual information in a small window approximating the exploration size of a finger (Loomis, Klatzky & Lederman, 1991). To fully compare AT and AV associations, taking differences in exploration styles into account, future studies can present visual information sequentially through a peek hole similar to that adopted by Loomis et al. (1991) or by Overvliet and Krampe (2018). It is likely that the development of AT correspondences involving different types of haptic exploration strategies, such as associating a /kiki/ sound and a rough texture such as sandpaper or abrasive sponges (Etzi, Spence, Zampini, & Gallace, 2016) and associating a /kiki/ sound and a 3D spiky shape, are likely to mature more quickly than the more abstract AT associations tested here. Furthermore, AT correspondences in other domains, such as those based on stimulus frequency (e.g., auditory amplitude modulation and tactile vibration frequency), might emerge sooner than AV correspondences in the same domain (e.g., amplitude modulation and visual spatial frequency). Such correspondences have been shown in adults (e.g., Guzman-Martinez, Ortega, Grabowecky, Mossbridge, & Suzuki, 2012) but not yet in children. The generalizability of our findings to other types of crossmodal correspondences remains to be tested.

Notwithstanding the inherent processing differences between vision and touch, our results are largely in line with previous literature showing relatively protracted development for AT integration compared with AV integration in simultaneity judgments (Stanley et al., 2019). Our work adds to this line of literature by showing that such developmental differences also apply to sound–shape correspondence, a type of multisensory processing that does not necessarily involve the integration of multisensory information as shown in previous studies.

The role of visual input supporting AT development

We found that AT association is strengthened if younger participants have prior experience with seeing the unique abstract shapes before tactile testing (Experiment 2–AVAT condition). Given that task instruction and block duration is the same regardless of whether AT or AV is tested first, the significant AT association after AV experience rules out alternative accounts such as 6- to 8-year-olds not understanding our tactile task or not being engaged enough to complete the experiment due to too many trials. In fact, better performance on the AT task when it followed rather than preceded the AV task suggests that performance is not limited by children's limited attention span or by task difficulty or fatigue.

Seeing the abstract shape first in the AVAT condition in Experiment 2 might improve the efficiency of tactile-to-visual translation feeding into visual shape representation. Previous research suggests that factors such as image vividness and image difficulty were correlated with haptic recognition performance for 2D outlines (Lederman, Klatzky, Chataway, & Summers, 1990); adult participants tend to take longer and are less accurate in recognizing 2D outlines when the object image is difficult to imagine or highly unfamiliar. By allowing children to see the visual shapes prior to AT testing (where visual information is not available such that there is no requirement of haptic–visual integration), it is possible that children become more familiar with the abstract shapes or build a prototypical category of these shapes, which may make visual imagery of these shapes from tactile inputs easier later on. This easier and improved mental imagery via touch after first having seen the shapes might explain the order effects observed in Experiment 2. This interpretation of our finding is in line with previous research showing reduced or at-chance tactile shape–sound correspondence in early-blind individuals exploring abstract shapes (Fryer et al., 2014; Hamilton-Fletcher et al., 2018; Sourav et al., 2019). Having first seen the shapes might have also affected the subsequent strategies children employed to explore the shapes. For example, visually comparing a pair of abstract shapes highlights the critical difference between them—contour spikiness. Having that knowledge is likely to prompt participants to use haptic exploration strategies that are more effective in revealing contour spikiness.

The manifestation of one type of multisensory perception can scaffold other types of multisensory perception. Rohlf, Li, Bruns, and Röder (2020) showed that although 5-year-olds exhibit AV integration in the ventriloquist effect (biased localization of auditory stimuli toward simultaneously presented but spatially discrepant visual stimuli), they do not recalibrate subsequent sensory experience based on the illusion, unlike older children and adults. This study suggests that multisensory integration precedes crossmodal recalibration—or, more broadly speaking, learning—during development. Contextualizing our study in this broader framework, the early emergence of sound–shape correspondence between a sound and a visual shape could be used to calibrate subsequent multisensory experiences such as supporting the emergence of sound–shape correspondence between a sound and a tactile shape. This possibility remains to be empirically tested by studies investigating various types of crossmodal correspondence and multisensory perception in the same individuals across life span development.

Conclusion

To our knowledge, this is the first study examining the development of correspondences between nonsense words and abstract *tactile* shapes in children for which little is known. This study is also the first to directly compare this development with the development of correspondences between nonsense words and abstract *visual* shapes, which has already been studied extensively. It is unlikely that low-level sensory mechanisms can explain these associations. The enhancement of AT association strength after AV exposure speaks against such a mechanism. Our results highlight that crossmodal correspondences between what we hear and what we touch can be modulated by visual experience using a typically developing sample, complementing previous research from early-blind individuals. The underlying behavioral and neural mechanisms by which select visual experience enhances haptic object recognition as well as the linking of AT features, and how these mechanisms allow for the development and decline of crossmodal correspondences across the life span, remain to be explored.

Acknowledgments

We thank graduate student Sarah Izen and undergraduate research assistants in our lab (Shan Luo, Victoria Sanchez, Tracy Tam, Alexandra Sabov, Brianna Leonardo, Xenia Leviyah, and Erinda Morina) for all their hard work and help with data collection and analysis. We thank Sylvia Guillory, Hayley Smith, Adam Jacobs, and several anonymous reviewers for helpful feedback on earlier versions of the manuscript. We also acknowledge our public outreach collaboration. This research was conducted in Living Laboratory at the Museum of Science, Boston. Finally, none of this work would have been possible without generous support from the UMass Boston Dean's Office (V.M.C.) and the UMass Boston Undergraduate Research Fund (A.S., B.L.).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jecp.2021.105167.

References

Alexander, J. M., Johnson, K. E., & Schreiber, J. B. (2002). Knowledge is not everything: Analysis of children's performance on a haptic comparison task. *Journal of Experimental Child Psychology, 82*, 341–366.

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.

Audacity Team. (2014). Audacity: Free audio editor and recorder (Version 2.0.0) [computer program]. Retrieved April 20, 2014, from http://audacity.sourceforge.net/

Bahrick, L. E., & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Developmental Psychology, 36*, 190–201.

Bahrick, L. E., & Lickliter, R. (2004). Infants' perception of rhythm and tempo in unimodal and multimodal stimulation: A developmental test of the intersensory redundancy hypothesis. *Cognitive, Affective, & Behavioral Neuroscience, 4*, 137–147.

Barutchu, A., Crewther, D. P., & Crewther, S. G. (2009). The race that precedes coactivation: Development of multisensory facilitation in children. *Developmental Science, 12*, 464–473.

Barutchu, A., Danaher, J., Crewther, S. G., Innes-Brown, H., Shivdasani, M. N., & Paolini, A. G. (2010). Audiovisual integration in noise by children and adults. *Journal of Experimental Child Psychology, 105*, 38–50.

Barutchu, A., Toohey, S., Shivdasani, M. N., Fifer, J. M., Crewther, S. G., Grayden, D. B., & Paolini, A. G. (2019). Multisensory perception and attention in school-age children. *Journal of Experimental Child Psychology, 180*, 141–155.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B, 57*, 289–300.

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

Bremner, A. J., Mareschal, D., Lloyd-Fox, S., & Spence, C. (2008). Spatial localization of touch in the first year of life: Early influence of a visual spatial code and the development of remapping across changes in limb position. *Journal of Experimental Psychology: General, 137*, 149–162.

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*, 26681.

Chen, Y.-C., Huang, P.-C., Woods, A., & Spence, C. (2019). I know that "Kiki" is angular: The metacognition underlying sound–shape correspondences. *Psychonomic Bulletin & Review, 26*, 261–268.

Chen, Y.-C., Shore, D. I., Lewis, T. L., & Maurer, D. (2016). The development of the perception of audiovisual simultaneity. *Journal of Experimental Child Psychology, 146*, 17–33.

Chow, H. M., & Ciaramitaro, V. M. (2019). What makes a shape "baba"? The shape features prioritized in sound–shape correspondence change with development. *Journal of Experimental Child Psychology, 179*, 73–89.

Diederich, A., & Colonius, H. (2004). Bimodal and trimodal multisensory enhancement: Effects of stimulus onset and intensity on reaction time. *Perception & Psychophysics, 66*, 1388–1404.

Doehrmann, O., & Naumer, M. J. (2008). Semantics and the multisensory brain: How meaning modulates processes of audio-visual integration. *Brain Research, 1242*, 136–150.

Dunn, W., Griffith, J., Morrison, M., Tanquary, J., Sabata, D., Victorson, D., ... Gershon, R. (2013). Somatosensation assessment using the NIH Toolbox. *Neurology, 80*, S41–S44.

Etzi, R., Spence, C., Zampini, M., & Gallace, A. (2016). When sandpaper is "kiki" and satin is "bouba": An exploration of the associations between words, emotional states, and the tactile attributes of everyday materials. *Multisensory Research, 29*, 133–155.

Fernández-Prieto, I., Navarra, J., & Pons, F. (2015). How big is this sound? Crossmodal association between pitch and size in infants. *Infant Behavior and Development, 38*, 77–81.

Fontana, F. (2013). Association of haptic trajectories to Takete and Maluma. In I. Oakley & S. Brewster (Eds.). *Haptic and audio interaction design* (Vol. 7989, pp. 60–68). Berlin: Springer.

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.

Fryer, L., Freeman, J., & Pring, L. (2014). Touching words is not enough: How visual experience influences haptic–auditory associations in the "Bouba–Kiki" effect. *Cognition, 132*, 164–173.

Gori, M., Del Viva, M., Sandini, G., & Burr, D. C. (2008). Young children do not integrate visual and haptic form information. *Current Biology, 18*, 694–698.

Gori, M., Sandini, G., & Burr, D. (2012). Development of visuo-auditory integration in space and time. *Frontiers in Integrative Neuroscience, 6*. https://doi.org/10.3389/fnint.2012.00077.

Guzman-Martinez, E., Ortega, L., Grabowecky, M., Mossbridge, J., & Suzuki, S. (2012). Interactive coding of visual spatial frequency and auditory amplitude-modulation rate. Current Biology, 22, 383–388.
Hamilton-Fletcher, G., Pisanski, K., Reby, D., Stefańczyk, M., Ward, J., & Sorokowska, A. (2018). The role of visual experience in the emergence of cross-modal correspondences. *Cognition, 175*, 114–121.
Kalagher, H., & Jones, S. S. (2011). Developmental change in young children's use of haptic information in a visual task: The role of hand movements. *Journal of Experimental Child Psychology, 108*, 293–307.
Köhler, W. (1929). *Gestalt psychology*. New York: Liveright.
Köhler, W. (1947). *Gestalt psychology* (2nd ed.). New York: Liveright.
Lederman, S. J., & Klatzky, R. L. (1987). Hand movements: A window into haptic object recognition. *Cognitive Psychology, 19*, 342–368.
Lederman, S. J., Klatzky, R. L., Chataway, C., & Summers, C. D. (1990). Visual mediation and the haptic recognition of two-dimensional pictures of common objects. *Attention, Perception, & Psychophysics, 47*, 54–64.
Lewkowicz, D. J. (1996). Perception of auditory-visual temporal synchrony in human infants. *Journal of Experimental Psychology: Human Perception and Performance, 22*, 1094–1106.
Loomis, J. M., Klatzky, R. L., & Lederman, S. J. (1991). Similarity of tactual and visual picture recognition with limited field of view. *Perception, 20*, 167–177.
Margiotoudi, K., & Pulvermuller, F. (2020). Action sound–shape congruencies explain sound symbolism. *Scientific Reports, 10*, 12706.
Maurer, D., Pathman, T., & Mondloch, C. J. (2006). The shape of boubas: Sound–shape correspondences in toddlers and adults. *Developmental Science, 9*, 316–322.
Meredith, M. A., & Stein, B. E. (1983). Interactions among converging sensory inputs in the superior colliculus. *Science, 221*, 389–391.
Nava, E., Grassi, M., & Turati, C. (2016). Audio-visual, visuo-tactile and audio-tactile correspondences in preschoolers. *Multisensory Research, 29*, 93–111.
Neil, P. A., Chee-Ruiter, C., Scheier, C., Lewkowicz, D. J., & Shimojo, S. (2006). Development of multisensory spatial integration and perception in humans. *Developmental Science, 9*, 454–464.
Overvliet, K. E., & Krampe, R. T. (2018). Haptic two-dimensional shape identification in children, adolescents, and young adults. *Journal of Experimental Child Psychology, 166*, 567–580.
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.
Petrini, K., Remark, A., Smith, L., & Nardini, M. (2014). When vision is not an option: Children's integration of auditory and haptic information is suboptimal. *Developmental Science, 17*, 376–387.
R Core Team (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
Rohlf, S., Li, L., Bruns, P., & Röder, B. (2020). Multisensory integration develops prior to crossmodal recalibration. *Current Biology, 30*, 1726–1732.e7.
Schwarzer, G., Küfer, I., & Wilkening, F. (1999). Learning categories by touch: On the development of holistic and analytic processing. *Memory & Cognition, 27*, 868–877.
Sourav, S., Kekunnaya, R., Shareef, I., Banerjee, S., Bottari, D., & Röder, B. (2019). A protracted sensitive period regulates the development of cross-modal sound–shape associations in humans. *Psychological Science, 30*, 1473–1482.
Spector, F., & Maurer, D. (2011). The colors of the alphabet: Naturally-biased associations between shape and color. *Journal of Experimental Psychology: Human Perception and Performance, 37*, 484–495.
Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics, 73*, 971–995.
Spence, C. (2013). Just how important is spatial coincidence to multisensory integration? Evaluating the spatial rule. *Annals of the New York Academy of Sciences, 1296*, 31–49.
Spence, C., & Sathian, K. (2020). Audiovisual crossmodal correspondences: Behavioural consequences and neural underpinnings. In K. Sathian & V. S. Ramachandran (Eds.), *Multisensory perception: From laboratory to clinic* (pp. 239–258). Cambridge, MA: Academic Press.
Stanley, B. M., Chen, Y.-C., Lewis, T. L., Maurer, D., & Shore, D. I. (2019). Developmental changes in the perception of audiotactile simultaneity. *Journal of Experimental Child Psychology, 183*, 208–221.
Taylor, S., McLean, B., Falkmer, T., Carey, L., Girdler, S., Elliott, C., & Blair, E. (2016). Does somatosensation change with age in children and adolescents? A systematic review. *Child: Care. Health and Development, 42*, 809–824.
Todd, J. W. (1912). *Archives of psychology: Reaction to multiple stimuli*. The Science Press.
Torchiano, M. (2017). effsize: Efficient effect size computation. R package Version 0.7.1. https://CRAN.R-project.org/package=effsize
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.
Vroomen, J., & De Gelder, B. (2004). Perceptual effects of cross-modal stimulation: Ventriloquism and the freezing phenomenon. In G. A. Calvert, C. Spence, & B. E. Stein (Eds.), *The handbook of multisensory processes* (pp. 141–150). Cambridge, MA: MIT Press.
Wang, C. A., Blohm, G., Huang, J., Boehnke, S. E., & Munoz, D. P. (2017). Multisensory integration in orienting behavior: Pupil size, microsaccades, and saccades. *Biological Psychology, 129*, 36–44.
Withagen, A., Kappers, A. M. L., Vervloed, M. P. J., Knoors, H., & Verhoeven, L. (2012). Haptic object matching by blind and sighted adults and children. *Acta Psychologica, 139*, 261–271.
Withagen, A., Kappers, A. M. L., Vervloed, M. P. J., Knoors, H., & Verhoeven, L. (2013). The use of exploratory procedures by blind and sighted adults and children. *Attention, Perception, & Psychophysics, 75*, 1451–1464.