Prelinguistic Infants Are Sensitive to Space-Pitch Associations Found Across Cultures

Sarah Dolscheid[1,2,3], Sabine Hunnius[4], Daniel Casasanto[5], and Asifa Majid[4,6,7]

[1]Neurobiology of Language Department, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands;
[2]International Max Planck Research School for Language Sciences, Nijmegen, The Netherlands;
[3]Department of Rehabilitation and Special Education, University of Cologne; [4]Donders Institute for Brain,
Cognition and Behaviour, Radboud University Nijmegen; [5]Department of Psychology, The University of Chicago;
[6]Language and Cognition Department, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands; and
[7]Center for Language Studies, Radboud University Nijmegen

Abstract

People often talk about musical pitch using spatial metaphors. In English, for instance, pitches can be "high" or "low" (i.e., height-pitch association), whereas in other languages, pitches are described as "thin" or "thick" (i.e., thickness-pitch association). According to results from psychophysical studies, metaphors in language can shape people's nonlinguistic space-pitch representations. But does language establish mappings between space and pitch in the first place, or does it only modify preexisting associations? To find out, we tested 4-month-old Dutch infants' sensitivity to height-pitch and thickness-pitch mappings using a preferential-looking paradigm. The infants looked significantly longer at cross-modally congruent stimuli for both space-pitch mappings, which indicates that infants are sensitive to these associations before language acquisition. The early presence of space-pitch mappings means that these associations do not originate from language. Instead, language builds on preexisting mappings, changing them gradually via competitive associative learning. Space-pitch mappings that are language-specific in adults develop from mappings that may be universal in infants.

Does a cake taste yellow? Or does a tone played by a trumpet sound scarlet? For some people they do. Yet synesthesia, a condition in which stimulation of one sensory modality induces systematic perceptual experiences in another modality, is relatively rare. Other types of cross-modal associations, however, can be found in nonsynesthetes as well, especially in the domain of musical pitch (Spence, 2011). Psychophysical studies have shown that adults and children without synesthesia associate higher pitches with sharper edges (Marks, 1987; Parise & Spence, 2009), lighter color (Hubbard, 1996; Marks, 1989; Melara, 1989), and positions higher in space (Ben-Artzi & Marks, 1995; Evans & Treisman, 2010).

Even infants seem to be sensitive to some of these associations (Haryu & Kajikawa, 2012; Jeschonek, Pauen,

& Babocsai, 2012; Wagner, Winner, Cicchetti, & Gardner, 1981; Walker et al., 2010). In a preferential-looking task, 3- to 4-month-olds preferred congruent trials in which visuospatial height and pitch corresponded (Walker et al., 2010). That is, infants looked longer at a ball moving upward when it was accompanied by a rising pitch than when it was accompanied by a falling pitch. Although the early presence of cross-modal associations has led some researchers to conclude that these mappings are probably

Corresponding Author:
Sarah Dolscheid, Department of Rehabilitation and Special Education, University of Cologne, Herbert-Lewin-Straße 10, D-50931 Cologne, Germany
E-mail: sarah_dolscheid@gmx.de

a

b
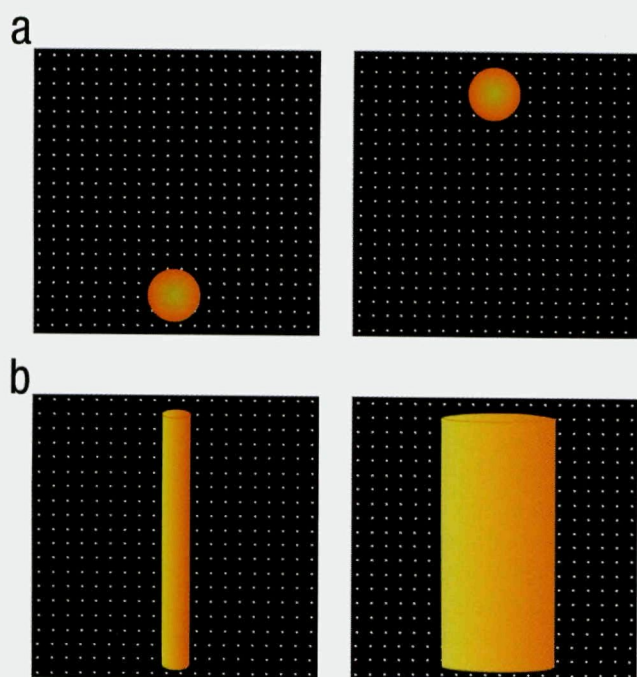
Fig. 1. Stills from animations in the (a) height-pitch task and (b) thickness-pitch task. The images, which are reproduced to scale, show the extremes of the ball's vertical trajectory (height-pitch task) and the tube's thickness (thickness-pitch task).

innately hardwired in the brain (Mondloch & Maurer, 2004), others point out that they change during the course of development (Marks, Hammeal, & Bornstein, 1987; Smith & Sera, 1992). Metaphorical language, in particular, is one factor that may alter cross-modal pitch associations (e.g., Martino & Marks, 1999).

Empirical support for the influence of language on space-pitch associations comes from cross-linguistic psychophysical studies. Whereas languages such as English and Dutch encode pitch in terms of height, other languages, such as Farsi, Turkish, and Zapotec (spoken in Mexico), use a thickness metaphor: High-frequency pitches are described as thin, and low-frequency pitches are described as thick (Shayan, Ozturk, & Sicoli, 2011). In a prior study, we (Dolscheid, Shayan, Majid, & Casasanto, 2013) tested whether these different linguistic metaphors influenced nonlinguistic space-pitch associations. Dutch and Farsi participants were asked to reproduce (i.e., sing back) musical pitches they heard in the presence of irrelevant spatial information (i.e., lines varying in either height or thickness). Dutch speakers' pitch estimates were modulated by spatial height, but not thickness. Conversely, Farsi speakers' pitch estimates were modulated by spatial thickness, but not height. In summary, nonlinguistic space-pitch associations followed language-specific metaphors, thus demonstrating that cross-modal pitch representations can be affected by language.

It is unclear, however, whether language establishes cross-modal mappings between space and pitch in the first place or simply modifies preexisting associations. Boroditsky (2000) posited that exposure to space-time metaphors in language leads to the construction of some nonlinguistic mappings between space and time; in principle, the same could hold for space and pitch. Alternatively, space-pitch metaphors in language could reflect earlier developmental cross-domain mappings, which could be innate or learned (on the basis of correlations between space and pitch in an infant's environment).

Infants seem to be sensitive to height-pitch mappings even prelinguistically (Walker et al., 2010). But are infants also sensitive to the thickness-pitch relationship, or is that association learned only on the basis of language input? It is possible that the height-pitch metaphor is privileged in language and cognition. Some researchers have speculated that the association of spatial height and pitch could be a consequence of the roughly linear place coding of pitches in the cochlea, from the apex to the base (cf. Pratt, 1930). In addition, other researchers have argued that the height-pitch metaphor is reflected in most languages (Parkinson, Kohler, Sievers, & Wheatley, 2012). Moreover, this association appears to be represented cognitively even when the corresponding linguistic metaphor is absent (Parkinson et al., 2012). In contrast, the evidence for the cross-cultural robustness of the thickness-pitch association is sparse (e.g., Shayan et al., 2011). Perhaps this association is learned only after a child has been exposed to thickness-pitch metaphors in language.

To determine the prelinguistic availability of space-pitch mappings, we tested 4-month-old Dutch infants using a preferential-looking paradigm. To investigate height-pitch correspondences, we followed the procedure of Walker et al. (2010). Infants watched a ball moving up and down a screen (see Fig. 1a), and the movement was accompanied by the sound of a sliding whistle. The whistle's fundamental frequency changed at a constant rate. In the congruent condition, the pitch rose and fell in accordance with the ball's movement. In the incongruent condition, the pitch rose and fell in opposition to the ball's movement.

We also tested 4-month-old infants in a thickness-pitch task analogous to the height-pitch task. Instead of balls moving up and down the screen, the stimuli were vertical tubes that varied in thickness, changing continuously between thin and thick (see Fig. 1b). In the congruent condition, pitch rose and fell in accordance with the tube's contraction and expansion; that is, the tube expanded when pitch fell, which is congruent with the thickness-pitch metaphor found in a number of languages. In the incongruent condition, the pitch rose and

fell in opposition to the tube's contraction and expansion (i.e., the tube expanded when the pitch rose).

If both height-pitch and thickness-pitch mappings are available to infants prelinguistically, infants should prefer congruent height-pitch and congruent thickness-pitch stimuli over incongruent ones. If, however, height-pitch and thickness-pitch relationships follow different developmental trajectories, and thickness mappings are acquired later, then prelinguistic infants should show a preference for congruent height-pitch stimuli, but not for congruent thickness-pitch stimuli.

Experiment

Method

Participants. Ten male and 10 female infants completed the height-pitch task (mean age = 129 days, range = 113–138 days). Another 7 infants were tested but were excluded because of fussiness (5 infants) or experimenter error (2 infants). A different set of 10 male and 10 female infants completed the thickness-pitch task (mean age = 127 days, range = 113–138 days). Eight additional infants were tested but were excluded because of technical problems (1 infant) or fussiness (7 infants). Most infants completed only this study; a few participated in another study as well (reported elsewhere), but this experiment was always administered first.

Materials and procedure. QuickTime animations were presented on a 102- × 76-cm LCD monitor (Sony, Tokyo, Japan) using Habit X (http://habit.cmb.ucdavis .edu) software. Animations appeared within a 67- × 67-cm screen area (25.6° × 25.6° of visual arc) and lasted a maximum of 60 s. Before each animation, a flashing light called infants' attention to the screen. Infants sat in a Maxi-Cosi infant seat (Dorel Juvenile, Columbus, IN) that was placed on a parent's lap. They viewed the animations from a distance of approximately 1.5 m. Infants' visual fixations were monitored and recorded on video. Animations were stopped if the infant looked elsewhere for a single period of 1 s or more. We used frame-by-frame coding of each digitized video (SuperCoder; Hollich, 2008) to determine the total time that the infant looked at the animation. Coding was performed blind to the experimental condition. Twenty-five percent of the data were double-coded by a second coder who was also blind to the condition.

In the height-pitch task, infants watched an orange ball, 10 cm (4°) in diameter, moving up and down a 50-cm vertical trajectory in front of a 20 × 20 grid of small, white dots on a black field (Fig. 1a). The ball moved at a constant speed of 20 cm per second and paused for 42 ms at each endpoint. Animations were accompanied by the sound of a sliding whistle (a sinusoidal tone). The fundamental frequency of the sound changed at a constant rate, between 300 and 1700 Hz over 2.5 s, which coincided with a single phase of the animation (e.g., the ball's movement from the bottom to the top of its path). The sound paused briefly when the ball was stationary at its lowest and highest points. The amplitude of the sound increased and then decreased between 47 and 84 dB within each phase of the animation, peaking when the sound reached 1000 Hz (i.e., the midpoint between the highest and lowest pitches). Amplitude thus changed about twice as fast as pitch to ensure that variation in perceived pitch was not confounded with variation in loudness.

In the thickness-pitch task, infants watched a vertical orange tube that varied in thickness,[1] changing continuously from thin to thick and then to thin again (see Fig. 1b). The animation was presented on a 20 × 20 grid of small, white dots on a black field, as in the height-pitch task. The tube was 60 cm long; its width ranged from 6 to 26 cm. It expanded at a constant speed of 8 cm per second and paused for 42 ms at each endpoint (i.e., its thickest and thinnest widths). Animations were accompanied by the sound of the sliding whistle used in the height-pitch task. The sound paused briefly when the tube was at its extremes. The parameters for pitch change and amplitude variation were identical to those in the height-pitch task.

For both tasks, infants viewed three congruent animations interleaved with three incongruent animations. Half the children watched a congruent animation first. Parents listened to music via headphones during the experiment so that they were unable to cue the infants inadvertently.

Results

The two observers agreed in their coding of each infant's looking times in the height-pitch task, ICC(3,1) (intraclass correlation coefficient) = .99, $F(29, 30) = 297.78$, $p < .001$, and thickness-pitch task, ICC(3,1) = .99, $F(29, 30) = 312.31$, $p < .001$.

According to Kolmogorov-Smirnov tests, looking times were normally distributed (all $ps > .05$). Looking times were compared using a 2 (spatial variation: height, thickness) × 2 (congruency: congruent, incongruent) mixed-factors analysis of variance. There was a main effect of congruency, $F(1, 38) = 8.53$, $p = .006$, $\eta_p^2 = .18$; infants looked longer at the congruent than at the incongruent stimuli. There was also a main effect of spatial variation, $F(1, 38) = 4.40$, $p = .04$, $\eta_p^2 = .10$. Infants looked longer at height stimuli than at thickness stimuli, which perhaps suggests that height was more salient. However, there was no interaction between spatial

variation and congruency, $F(1, 38) = 0.03$, $p = .86$, $\eta_p^2 = .00$, and thus no indication that looking-time congruency effects differed between tasks. The reduction in looking time on incongruent trials was comparable in the two tasks (i.e., 18% for the height-pitch task and 20% for the thickness-pitch task).

We also examined the effect of congruency in each task separately. Infants looked longer at the congruent animations in both the height-pitch task, $t(19) = 1.99$, $p = .06$, $d = 0.45$ (congruent trials: $M = 31.7$ s, $SD = 11.4$; incongruent trials: $M = 26.1$ s, $SD = 13.3$) and the thickness-pitch task, $t(19) = 2.19$, $p = .04$, $d = 0.43$ (congruent trials: $M = 24.4$ s, $SD = 11.8$; incongruent trials: $M = 19.4$ s, $SD = 11.5$), although the p value for the height-pitch task is above the conventional level of significance.

General Discussion

Our results demonstrate that prelinguistic infants are sensitive to at least two different types of space-pitch correspondence. Dutch 4-month-olds looked longer at audiovisual stimuli congruent, rather than incongruent, with height-pitch metaphors found in languages such as English and Dutch (these results are consistent with those of Walker et al., 2010). Four-month-olds also looked longer at stimuli congruent, rather than incongruent, with thickness-pitch metaphors found in languages such as Turkish and Farsi. Infants' tendency to look longer at congruent stimuli was similar in the height and thickness tasks, which suggests a comparable starting point for height-pitch and thickness-pitch mappings.

Are cross-modal pitch mappings therefore innate? On the basis of the current evidence, we can conclude only that these associations are present early in infancy. By the age of 4 months, however, infants may have encountered enough relevant environmental co-occurrences to have learned these mappings (see Lewkowicz, 2011). For example, people with bigger ("thicker") bodies tend to have lower voices, bigger bells produce lower tones, and so forth. Height-pitch mappings may also be grounded in bodily experience, because people's larynxes rise when they produce higher pitches and descend when they produce lower pitches (e.g., Miller, 1986).

Our findings show that infants are sensitive to space-pitch correspondences before they know linguistic space-pitch metaphors. Although language does not seem to create these mappings, linguistic metaphors could still influence the structure and content of preexisting mental representations via simple learning mechanisms. In the course of language acquisition, the relative strengths of different space-pitch mappings could be adjusted according to specific metaphors that children acquire (e.g., Casasanto, 2008, 2010). This process may parallel the acquisition of other linguistic

systems. In speech perception, for instance, infants start out as universal listeners but over time become language-specific listeners (e.g., Werker & Tees, 1984). Similar observations have been made for some semantic distinctions (e.g., Hespos & Spelke, 2004).

Regarding space-pitch metaphors, speaking a language with a height-pitch mapping, such as Dutch, could strengthen the height-pitch mapping at the expense of the thickness-pitch mapping, and the reverse may be true for a language with a thickness-pitch mapping, such as Farsi. Evidence in support of this competitive associative-learning account is provided by linguistic training experiments. For example, adult Dutch speakers, after being trained to use thickness metaphors to describe pitch relationships (as in Farsi), demonstrated nonlinguistic thickness-pitch mappings very similar to those of Farsi speakers (Dolscheid et al., 2013). These training studies demonstrate a causal role for language in strengthening some nonlinguistic mappings over others. At the same time, these results provide evidence that speakers can quite easily be retrained to use nonnative pitch representations; in contrast, it is difficult to retrain speakers to distinguish phonological contrasts not present in their native language (perceptual narrowing; e.g., Werker & Tees, 1984).

Conclusions

The finding that both height-pitch and thickness-pitch mappings can be observed in infants as young as 4 months old constrains theorizing about the role of language in shaping nonlinguistic mental representations of pitch. Our data show that space-pitch associations are present before language acquisition, which suggests that language is unlikely to create cross-modal mappings between space and pitch, even if language seems to create new mappings in other domains (Gentner, 2002). People who use different spatial metaphors for pitch in their native languages come to think about pitch differently not because language instills in them one cross-modal mapping instead of the other, but because language strengthens one preexisting mapping at the expense of the other.

Author Contributions

Note

1. Even though movement was restricted to the horizontal plane, which is characteristic for thickness, the stimuli displayed a concomitant difference in overall size. It is interesting in this regard that some languages use a size-pitch metaphor (Eitan & Timmers, 2010). In our study, nothing rested on distinguishing between thickness and size; the ability to make the cross-modal mapping to pitch would have been necessary for both spatial parameters.

References

Ben-Artzi, E., & Marks, L. E. (1995). Visual-auditory interaction in speeded classification: Role of stimulus difference. *Perception & Psychophysics, 57*, 1151–1162.

Boroditsky, L. (2000). Metaphoric structuring: Understanding time through spatial metaphors. *Cognition, 75*, 1–28.

Casasanto, D. (2008). Who is afraid of the big bad Whorf? Crosslinguistic differences in temporal language and thought. *Language Learning, 58*, 63–79.

Casasanto, D. (2010). Space for thinking. In V. Evans & P. Chilton (Eds.), *Language, cognition and space: The state of the art and new directions* (pp. 453–478). London, England: Equinox.

Dolscheid, S., Hunnius, S., Casasanto, D., & Majid, A. (2012). The sound of thickness: prelinguistic infants' associations of space and pitch. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Building bridges across cognitive sciences around the world: Proceedings of the 34th Annual Meeting of the Cognitive Science Society* (pp. 306–311). Retrieved from http://mindmodeling.org/cogsci2012/papers/0065/paper0065.pdf

Dolscheid, S., Shayan, S., Majid, A., & Casasanto, D. (2013). The thickness of musical pitch: Psychophysical evidence for linguistic relativity. *Psychological Science, 24*, 613–621. doi:10.1177/0956797612457374

Eitan, Z., & Timmers, R. (2010). Beethoven's last piano sonata and those who follow crocodiles: Cross-domain mappings of auditory pitch in a musical context. *Cognition, 114*, 405–422.

Evans, K. K., & Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features. *Journal of Vision, 10*(1), Article 6. Retrieved from http://www.journalofvision.org/content/10/1/6.full

Gentner, D. (2002). Analogy in scientific discovery: The case of Johannes Kepler. In L. Magnani & N. J. Nersessian (Eds.), *Model-based reasoning: Science, technology, values* (pp. 21–39). New York, NY: Kluwer Academic.

Haryu, E., & Kajikawa, S. (2012). Are higher-frequency sounds brighter in color and smaller in size? Auditory–visual correspondences in 10-month-old infants. *Infant Behavior & Development, 35*, 727–732.

Hespos, S. J., & Spelke, E. S. (2004). Conceptual precursors to language. *Nature, 430*, 453–456.

Hollich, G. (2008). Supercoder: A program for coding preferential looking (Version 1.7.1) [Computer software]. West Lafayette, IN: Purdue University.

Hubbard, T. L. (1996). Synesthesia-like mappings of lightness, pitch, and melodic interval. *The American Journal of Psychology, 109*, 219–238.

Jeschonek, S., Pauen, S., & Babocsai, L. (2012). Cross-modal mapping of visual and acoustic displays in infants: The effect of dynamic and static components. *European Journal of Developmental Psychology, 10*, 337–358.

Lewkowicz, D. J. (2011). The biological implausibility of the nature–nurture dichotomy and what it means for the study of infancy. *Infancy, 16*, 331–367.

Marks, L. E. (1987). On cross-modal similarity: Auditory–visual interactions in speeded discrimination. *Journal of Experimental Psychology: Human Perception and Performance, 13*, 384–394.

Marks, L. E. (1989). On cross-modal similarity: The perceptual structure of pitch, loudness, and brightness. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 586–602.

Marks, L. E., Hammeal, R. J., & Bornstein, M. H. (1987). Perceiving similarity and comprehending metaphor. *Monographs of the Society for Research in Child Development, 52*(1).

Martino, G., & Marks, L. E. (1999). Perceptual and linguistic interactions in speeded classification: Tests of the semantic coding hypothesis. *Perception, 28*, 903–923.

Melara, R. D. (1989). Dimensional interaction between color and pitch. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 69–79.

Miller, R. (1986). *The structure of singing: System and art in vocal technique*. New York, NY: Schirmer Books.

Mondloch, C. J., & Maurer, D. (2004). Do small white balls squeak? Pitch-object correspondences in your children. *Cognitive, Affective, & Behavioral Neuroscience, 4*, 133–136.

Parise, C., & Spence, C. (2009). "When birds of a feather flock together": Synesthetic correspondences modulate audiovisual integration in non-synesthetes. *PLoS ONE, 4*(5), Article e5664. Retrieved from http://dx.plos.org/10.1371/journal.pone.0005664

Parkinson, C., Kohler, P. J., Sievers, B., & Wheatley, T. (2012). Associations between auditory pitch and visual elevation do not depend on language: Evidence from a remote population. *Perception, 41*, 854–861.

Pratt, C. C. (1930). The spatial character of high and low tones. *Journal of Experimental Psychology, 13*, 278–285.

Shayan, S., Ozturk, O., & Sicoli, M. (2011). Thickness of pitch: Cross-modal metaphors in Farsi, Turkish and Zapotec. *The Senses & Society, 6*, 96–105.

Smith, L. B., & Sera, M. D. (1992). A developmental analysis of the polar structure of dimensions. *Cognitive Psychology, 24*, 99–142.

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics, 73*, 971–995.

Wagner, S., Winner, E., Cicchetti, D., & Gardner, H. (1981). "Metaphorical" mapping in human infants. *Child Development, 52*, 728–731.

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.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior & Development, 7*, 49–63.