Infant Behavior & Development 38 (2015) 77–81

Contents lists available at ScienceDirect

Infant Behavior and Development

Brief Report

How big is this sound? Crossmodal association between pitch and size in infants

Irune Fernández-Prieto[a,b,c,*], Jordi Navarra[c], Ferran Pons[a,b]

[a] Institute for Brain, Cognition and Behaviour (IR3C), Barcelona, Spain
[b] Departament de Psicologia Bàsica, Facultat de Psicologia, Universitat de Barcelona, Barcelona, Spain
[c] Fundació Sant Joan de Déu, Parc Sanitari Sant Joan de Déu – Hospital Sant Joan de Déu, Esplugues de Llobregat, Barcelona, Spain

ARTICLE INFO

Article history:
Received 7 August 2014
Received in revised form 6 November 2014
Accepted 18 December 2014
Available online 21 January 2015

Keywords:
Auditory pitch
Size
Infant perception
Crossmodal correspondence
Preferential looking

ABSTRACT

We examined 4- and 6-month-old infants' sensitivity to the perceptual association between pitch and object size. Crossmodal correspondence effects were observed in 6-month-old infants but not in younger infants, suggesting that experience and/or further maturation is needed to fully develop this crossmodal association.

High-frequency sounds are associated crossmodaly with high spatial positions, as well as with bright, small, and sharp objects. Conversely, low frequency sounds are associated with low spatial positions, and with dark, big and rounded objects (see Bien, ten Oever, Goebel, & Sack, 2012; Gallace & Spence, 2006; Marks, 1974; Melara & O'Brien, 1987; Rusconi, Kwan, Giordano, Umilta, & Butterworth, 2006; see also Spence, 2011, for a review).

A relevant question from a developmental point of view concerns whether these possible audiovisual correspondences have to be learned from experience or else they are already present from very early on, even before language is acquired. Some studies suggest that crossmodal correspondences are rooted in language (see Marks, 1984; Marks, Hammeal, & Bornstein, 1987; Smith & Sera, 1992). These associative links would, according to this hypothesis, emerge after the acquisition of language (that is, after specific space-related lexical items are acquired). In contrast, other studies indicate that language may not be essential for the arising of such crossmodal links. For example, Parkinson, Kohler, Sievers, and Wheatley (2012) explored pitch classification in adults from a hill tribe in Cambodia, whose language does not use spatial-related labels to refer to pitch (as it happens in English, and the vast majority of languages). The perceptual association between pitch and vertical position was present in members of this tribe, suggesting that the presence of space-based terms to describe pitch may not be necessary for this perceptual association to appear (see also Martino & Marks, 1999).

Regarding the development of these crossmodal associations, Walker and colleagues (2010) found evidence of a cross-modal link between dynamic sounds and visuospatial height (as well as between pitch and visual sharpness) in 3- to 4-month-old infants. Infants looked longer at the visual stimuli in the crossmodaly congruent conditions (e.g., an ascending

* Corresponding author at: Irune Fernández-Prieto, Parc Sanitari de Sant Joan de Déu, Fundació Sant Joan de Déu, c/ Santa Rosa 39-57, Edifici Docent, 4th floor, Esplugues de Llobregat, Barcelona 08950, Spain. Tel.: +34 936009751; fax: +34 936009771.
E-mail address: irunefernandez@gmail.com (I. Fernández-Prieto).

http://dx.doi.org/10.1016/j.infbeh.2014.12.008
0163-6383/

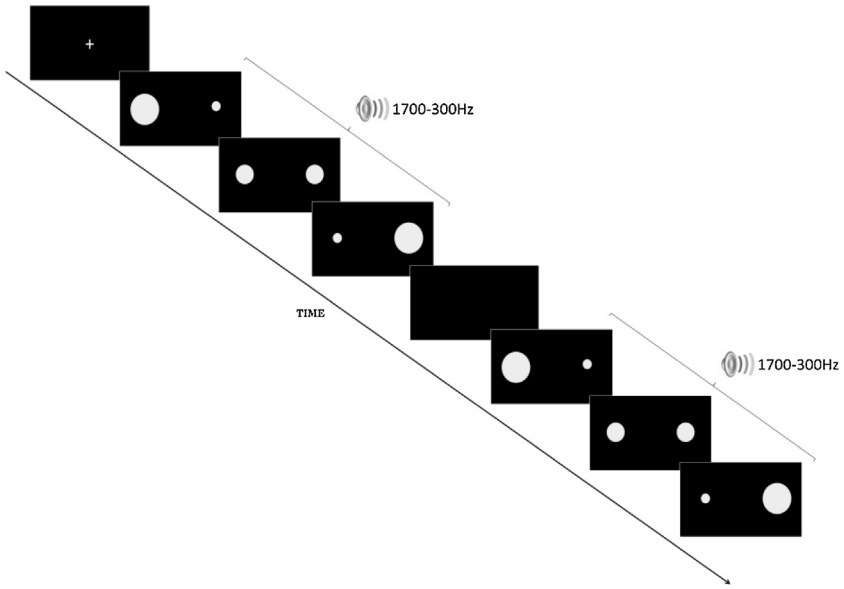

Fig. 1. Example of a sequence of events of two trial animations used in the experiment. The size of the left and the right ball, which were presented in combination with a descending frequency sweep, decreased and increased, respectively.

frequency sound combined with a visual stimulus moving toward the upper part of the screen) than in the incongruent conditions. Although other authors have failed to replicate these effects (see Lewkowicz & Minar, 2014), Dolscheid, Hunnius, Casasanto, and Majid (2014) have recently found converging evidence of a crossmodal association between ascending or descending sweeps (low-to-high or high-to-low in frequency) and spatial elevation in 3- to 4-month-old infants, and also between pitch and width. Infants preferred to look at the congruent condition (e.g., a narrow visual stimulus combined with an ascending frequency sweep) than at the incongruent one. In another recent study, Peña, Mehler, and Nespor (2011) also demonstrated that 4-month-old infants are able to perceptually associate 'high-frontal' (/i/and/e/) and 'low-posterior' (/a/and/o/) vowels with small and large geometric shapes, respectively.

Other perceptual associations between visual and acoustic features, such as size or brightness and pitch, have also been observed in 30–36 month-old children (Mondloch & Maurer, 2004). Haryu and Kajikawa (2012) recently reported that 10-month-old infants associate high/low frequencies with bright/dark objects, respectively. However, no signs of a perceptual relation between pitch and visual size were observed at this age.

In the current study, we further explored the association between pitch and the size of visual stimuli in young infants. In contrast with Peña et al.'s study (2011), we used auditory stimuli (pure-tone frequency sweeps) that bore no relation to human language. This allowed us to address possible perceptual links between visual size and pitch beyond possible "speech-based" (and perhaps more direct) associations (e.g., between different vowels and the size of the mouth aperture). Moreover, in contrast with Haryu and Kajikawa's (2012) and Mondloch & Maurer's (2004) studies, dynamic sounds (frequency sweeps) were used instead of low or high frequency flat tones, thus covering a larger spectrum of sound frequencies that ranged between 300 Hz and 1700 Hz (see Dolscheid et al., 2014; Walker et al., 2010). Four- and 6-month-old infants were tested to examine whether this correspondence is already present at these early ages and, therefore, before the acquisition of spatial/abstract lexical items or not (see Swingley, 2009, for a review).

Eighteen full-term 4-month-old infants (9 females; $M = 129$ days, $SD = 5$ days) and 18 full-term 6-month-old infants participated in the study (7 females; $M = 189$ days, $SD = 7$ days). Sixteen additional infants were tested but not included in the final sample because of crying or fussiness (eight 4-month-old and seven 6-month-old infants) and experimental error (one 4-month-old infant). Participants were recruited at the maternity unit at Hospital Sant Joan de Déu. Parental consent was obtained before running the experiment. During the experiment, the infant sat in a high chair, watching a series of audiovisual animations on a 17-inch TFT monitor. The total amount of time that an infant looked directly to a specific area of the screen was determined by a Tobii T120 Eye-tracker, which was used to collect and store eye-tracking data in an Intel Core i5 computer. Auditory stimuli were delivered from two loudspeakers located at each side of the computer screen.

The experiment contained 4 blocks, randomly presented, each one including 6 trials (24 trials in total). Each block began with an attention getter – a white cross – at the center of the screen, followed by 6 identical trials of 2500 ms of duration. Each trial consisted on two yellow balls appearing simultaneously at the 2 sides (left and right) of the screen, at 15.2 cm from the center (see Fig. 1.). While one of the balls increased in size (from a diameter of 3 cm to a diameter of 13.5 cm), the other one decreased (from a diameter of 13.5 cm to a diameter of 3 cm), at a constant speed of 4.2 cm per second. The presentation of these dynamic visual stimuli was accompanied by an auditory presentation of an ascending (300–1700 Hz)