View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by University of Birmingham Research Portal

UNIVERSITY OF BIRMINGHAM

Research at Birmingham

Synaesthetic interactions across vision and audition

Krugliak, Alexandra; Noppeney, Uta

DOI:
10.1016/j.neuropsychologia.2015.09.027

License:
Creative Commons: Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)

Document Version
Peer reviewed version

Citation for published version (Harvard):
Krugliak, A & Noppeney, U 2016, 'Synaesthetic interactions across vision and audition', Neuropsychologia, vol. 88, pp. 65-73. https://doi.org/10.1016/j.neuropsychologia.2015.09.027

Link to publication on Research at Birmingham portal

Publisher Rights Statement:
Checked 22/07/2016

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of ‘fair dealing’ under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

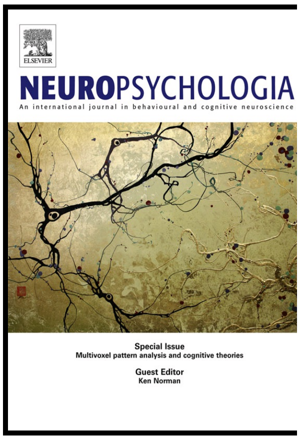
Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 01. Feb. 2019

Synesthetic interactions across vision and audition

Alexandra Krugliak, Uta Noppeney


www.elsevier.com/locate/neuropsychologia

PII: S0028-3932(15)30170-6
DOI: http://dx.doi.org/10.1016/j.neuropsychologia.2015.09.027
Reference: NSY5739

To appear in: *Neuropsychologia*

Received date: 22 April 2015
Revised date: 14 August 2015
Accepted date: 21 September 2015

Cite this article as: Alexandra Krugliak and Uta Noppeney, Syntheti interactions across vision and audition, *Neuropsychologi* http://dx.doi.org/10.1016/j.neuropsychologia.2015.09.027

Synesthetic interactions across vision and audition

Alexandra Krugliak, Uta Noppeney

Synesthetic interactions across vision and audition

Alexandra Krugliak, Uta Noppeney

Authors:

Alexandra Krugliak
Email: AXK263@bham.ac.uk
Centre for Computational Neuroscience and Cognitive Robotics
University of Birmingham
Edgbaston
Birmingham
B15 2TT
UK

Prof. Uta Noppeney
Email: U.Noppeney@bham.ac.uk
Centre for Computational Neuroscience and Cognitive Robotics
University of Birmingham
Edgbaston
Birmingham
B15 2TT
UK

Correspondence to:

Alexandra Krugliak
Email: AXK263@bham.ac.uk
Centre for Computational Neuroscience and Cognitive Robotics
University of Birmingham
Edgbaston
Birmingham
B15 2TT
UK

Correspondence:
Alexandra Krugliak
Email: AXK263@bham.ac.uk
Centre for Computational Neuroscience and Cognitive Robotics
University of Birmingham
Edgbaston
Birmingham
B15 2TT
UK

Abstract

In everyday life our senses are exposed to a constant influx of sensory signals. The brain binds signals into a coherent percept based on temporal, spatial or semantic correspondences. In addition, synesthetic correspondences may form important cues for multisensory binding. This study focussed on the synesthetic correspondences between auditory pitch and visual size. While high pitch has been associated with small objects in static contexts, recent research has surprisingly found that increasing size is linked with rising pitch.

The current study presented participants with small/large visual circles/discs together with high/low pitched pure tones in an intersensory selective attention paradigm. Whilst fixating a central cross participants discriminated between small and large visual size in the visual modality or between high and low pitch in the auditory modality. Across a series of five experiments, we observed convergent evidence that participants associated small visual size with low pitch and large visual size with high pitch. In other words, we observed the pitch-size mapping that has previously been observed only for dynamic contexts. We suggest that these contradictory findings may emerge because participants can interpret visual size as an index of permanent object size or distance (e.g. in motion) from the observer. Moreover, the pitch-size mapping may depend not only on relative but also on the absolute levels of pitch and size of the presented stimuli.

Keywords Multisensory Integration; Crossmodal Correspondence; Synesthesia; Auditory Pitch; Visual Size

1. Introduction

In daily life our brains are bombarded with myriad of signals perceived through different sensory modalities. Signals originating from a common event need to be integrated into one coherent percept and separated from other signals.

Temporal, spatial and semantic congruency are important cues that inform the brain whether signals originate from a common source and should be integrated
(Adam & Noppeney, 2010; van Atteveld, Formisano, Goebel, & Bloemert, 2004; van Atteveldt et al., 2007; Donohue, Roberts, Grent-'t-Jong, & Woldorff, 2011; Laurienti, Kraft, Maldjian, Burdette, & Wallace, 2004; Lee & Noppeney, 2014; Lewis & Noppeney, 2010; Macaluso & Driver, 2005; Lee & Noppeney, 2011; Vroomen & Keetels, 2010; Wallace, Wilkinson, & Stein, 1996; Wallace et al., 2004). In addition to these classical congruency cues more abstract feature correspondences can also influence the binding of signals from multiple sensory modalities.

The most pronounced examples are synesthetic experiences binding letters with colours or colours with sounds (Brang, Rouw, Ramachandran, & Coulson, 2011; Rich, Bradshaw & Mattingley, 2005). Yet, even in non-synesthetic individuals perceptual experiences and decisions are influenced by a wide range of multisensory metaphoric mappings including frequency-size (Antovic, 2009; Bien, ten Oever, Goebel, & Sack, 2012; Eitan, Schupak, Gotler, & Marks, 2014; Evan & Treisman, 2010; Gallace & Spence, 2006; Marks, Hammeal, & Bornstein, 1987; Mondloch & Maurer, 2004; Parise & Spence, 2009; Parise & Spence, 2012), dynamic pitch–dynamic size (Eitan et al, 2014; Fernandez-Prieto, Navarra & Pons, 2015; Kim & Iwamiya, 2008), and dynamic pitch–directional motion (Sadahiani, Maier & Noppeney, 2009) (for reviews see Marks, 2004; Spence, 2011; Spence & Deroy, 2013). For example, human observers perceive bright objects as louder than dark objects (loudness-brightness: Marks, 1987). They also tend to associate high-pitch sounds predominantly with visual objects at higher elevation (frequency-elevation: Ben-Artzi & Marks, 1995; Bernstein & Edelstein, 1971; Evans & Treisman, 2010; Melara & O'Brien, 1987; Patching & Quinlan, 2002).

Multiple mechanisms have been proposed to mediate metaphoric relationships. One account posits that metaphoric mappings are mediated via shared semantics or language. For instance, pitch is referred to by words such as 'high' or 'low'. Moreover, musical notation relies on spatial concepts (Martino & Marks, 1999; Ashley, 2004). Hence, interactions between pitch in the auditory sense and elevation in the visual sense may be mediated via a common conceptual reference frame. Alternatively, seemingly arbitrary metaphoric mappings may in fact be grounded in natural environmental statistics. In line with this conjecture, a recent elegant study by Parise, Knorre, and Ernst (2014) revealed that the mapping between frequency and elevation is grounded in auditory scene statistics where high-frequency sounds tend to originate from elevated sources. Moreover, the filtering characteristics of the outer ear also contributed to the mapping between elevation and frequency.

In a similar vein, the metaphoric relationship between auditory frequency and visual object size has been proposed to emerge from the fact that the frequency of sounds made by animals or musical instruments depends on the size of the resonator (von Kriegstein, Smith, Patterson, Ives, & Griffiths, 2007). In other words, high-pitched sounds should be associated with small objects and low-pitched sounds with large objects (Bien et al., 2012; Eitan et al, 2014; Evan & Treisman, 2010; Gallace & Spence, 2006; Marks et al., 1987; Mondloch & Maurer, 2004; Parise & Spence, 2009; Parise & Spence, 2012; except: Antovic, 2009).

While accumulating evidence associates high-frequency sounds with small objects and vice versa in a static context, controversial evidence has been provided for dynamic contexts.

Here, ascending pitch has surprisingly been associated with growing size (Eitan et al, 2014). Amongst other mechanisms, the authors attributed this opposite pattern for dynamic stimuli to the Doppler Effect whereby an approaching object induces a change in pitch. This experiment suggests an ambivalent association between pitch and size in our natural dynamic world. In dynamic contexts, the brain would need to dissociate whether the size as estimated from a retinotopic representation derives predominantly from the constant size of the object in the natural world or its distance from the observer. This more complex relationship between constant and dynamic size-pitch relationship may explain why the correspondence between pitch and size develops relatively late in life (Fernandez-Prieto et al., 2015; Marks et al., 1987; Mondloch & Maurer, 2004) and has been found only inconsistently (Haryn & Kajikawa, 2012; Mondloch & Maurer, 2004).

This study revisits the pitch-size relationship in a static context. Participants were presented with large or small circles/discs in synchrony with high- or low-frequency sounds in an auditory or visual selective attention paradigm (Bernstein & Edelstein, 1971). In the visual modality, they discriminated between large and small visual size. In the auditory modality, they discriminated between high- and low-pitched tones. As luminance may be a confounding factor when varying the size of a visual stimulus, the visual discs were either brighter or darker than the background colour. Likewise, loudness and sound amplitude can be potential confounds that we evaluated by equating the sounds either with respect to their physical sound amplitude or their perceptual loudness.

2. Experiment 1 & 2

2.1 Methods

2.1.1 Participants

After giving written informed consent, 16 participants (12 female, mean age: 24 years) took part in Experiment 1 and 10 participants (4 female, mean age: 23 years) in Experiment 2. Each had normal or corrected-to-normal vision, reported normal hearing, and had no history of neurological or psychiatric illness. The study was approved by the local research and ethics committee.

2.1.2 Stimuli

Visual stimuli were either circles (thickness: 0.5° visual angle) or discs. The radius of both circles and discs was either 2.8° or 7.7° visual angle. Experiment 1 presented circles or discs in lighter grey (mean luminance: 50.08 cd/m^2) than the grey shade of the background (mean luminance: 33.58 cd/m^2). Experiment 2 presented circles or discs in darker grey (mean luminance: 33.58 cd/m^2) than the grey shade of the background (mean luminance: 50.08 cd/m^2). The comparison between Experiment 1 and 2 allows us to assess confounding effects of luminance variation on the pitch-size association. This is important, because previous studies have demonstrated that pitch is not only associated with size but also with brightness (Marks, 1987; Marks et al., 1987). Yet, overall brightness differs between (i) circles and discs and (ii) in particular discs of different sizes.

Auditory stimuli were pure tones of 120 ms duration with linear onset and offset ramps of 1 ms to avoid auditory clicks (sampling rate 44100 Hz). The frequency was either 1250 Hz (low pitch) or 3000 Hz (high pitch).

2.1.3 Experimental design

The 3 x 2 factorial design manipulated: (i) visual stimuli (circles or discs), (ii) task-relevant modality (respond to the auditory or to the visual stimuli), and (iii) mapping (mapping 1: low pitch, large size and high pitch, small size; mapping 2: low pitch, small size and high pitch, large size).

On each trial participants were presented with an audiovisual stimulus (120 ms duration, SOA 1500 ms) defined by pitch (high, low) and size (large, small). Thus, four audiovisual stimulus combinations were presented with equal probability: low pitch/large visual size, low pitch/small visual size, high pitch/large visual size and high pitch/small visual size. We will refer to the stimulus combinations low/large and high/small as mapping 1 and to the stimulus combinations low/small and high/large as mapping 2 (Figure 1*).* In Experiment 1, the visual disc was brighter than the background. In Experiment 2, the visual disc was darker than the background. In a selective attention paradigm, participants performed a two-choice discrimination task that focussed either on the auditory frequency or the visual size dimension. Participants discriminated between small and large size in the visual task or high and low pitch in the auditory task as fast and accurately as possible. Further, they were instructed to fixate a central fixation cross throughout the entire experiment.

The main experiment included two runs presenting ‘circles’ or ‘discs’ respectively as visual stimuli. Each run consisted of 12 auditory and 12 visual attention task blocks that were presented in permuted order to facilitate interference effects. The task-relevant sensory modality was indicated at the beginning of each block. Within each block each of the four possible audiovisual stimulus combinations were presented twice in random order. The order of auditory/visual tasks and circle/disc runs was counterbalanced across participants. The start of each block was initiated by button press in order to allow participants to switch between the different response-mappings for the auditory and visual task.

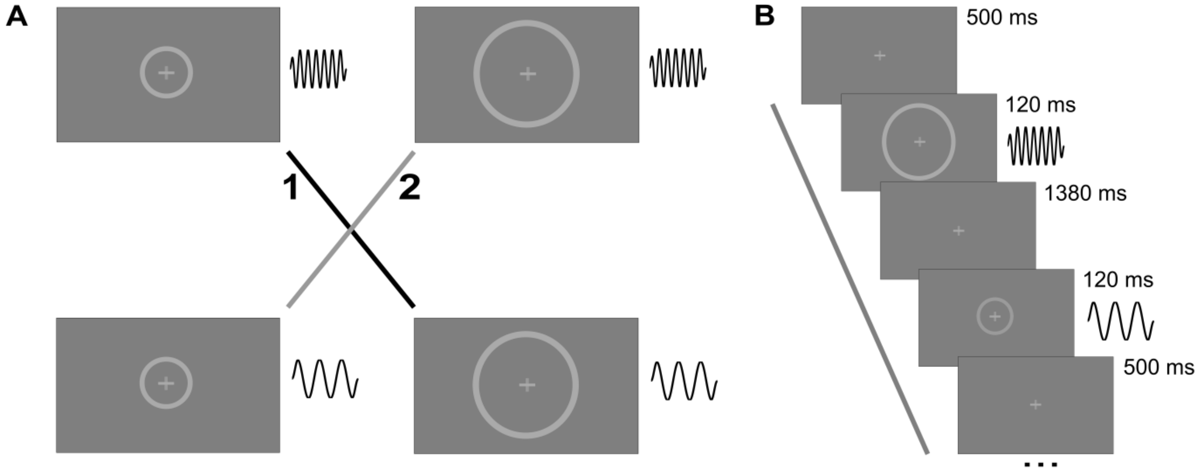


Figure 1*. *A. Experimental design: Each experiment compared two mappings: mapping 1: high pitch with small size and low pitch with large size; mapping 2: high pitch with large size and low pitch with small size. ***B.*** Example trial: On each trial participants were presented with a visual circle (or disc) and an auditory pure tone. In the auditory task, they discriminated between high and low pitched tones. In the visual task, they discriminated between small and large sized visual stimuli.

Responses were given via four different buttons: ‘A’, ‘D’, ‘J’ and ‘K on a conventional keyboard. The buttons were chosen to ensure that participants used different hands to respond during the auditory and visual tasks in order to avoid interference and transference effects at the response level. The mapping of hands and response buttons was counterbalanced across participants resulting in eight response mapping options:

1. auditory left hand and visual right hand: (i) ‘A’ = low (auditory), ‘D’ = high (auditory), ‘J’ = small (visual), ‘K’ = large (visual), (ii) ‘A’ = low (auditory), ‘D’ = high (auditory), ‘J’ = large (visual), ‘K’ = small (visual), (iii) ‘A’ = high (auditory), ‘D’ = low (auditory), ‘J’ = small (visual), ‘K’ = large (visual), (iv) ‘A’ = high (auditory), ‘D’ = low (auditory), ‘J’ = large (visual), ‘K’ = small (visual)
2. auditory right hand and visual left hand: (i) ‘J’ = low (auditory), ‘K’ = high (auditory), ‘A’ = small (visual), ‘D’ = large (visual), (ii) ‘J’ = low (auditory), ‘K’ = high (auditory), ‘A’ = large (visual), ‘D’ = small (visual), (iii) ‘J’ = high (auditory), ‘K’ = low (auditory), ‘A’ = small (visual), ‘D’ = large (visual), (iv) ‘J’ = high (auditory), ‘K’ = low (auditory), ‘A’ = large (visual), ‘D’ = small (visual)

2.1.4 Apparatus

The experiment was conducted in a dimly lit experimental room. Constant viewing distance was ensured by stabilizing the participant’s head on a chin rest at a distance of 50 cm from a LED monitor (1920 × 1080 resolution, 60 Hz refresh rate, iiyama Proline, Japan). Auditory stimuli were presented through headphones (Sennheiser HD 555MR, Germany) at approximately 75 dB SPL. Experimental sessions were presented using Cogent 2000 v1.25 (developed by the Cogent 2000 team at the FIL and the ICN and Cogent Graphics developed by John Romaya at the LON at the Wellcome Department of Imaging Neuroscience, UCL, London, UK; http://www.vislab.ucl.ac.uk/cogent.php) running under MATLAB (Mathworks Inc., Natick, MA, USA) on a Windows PC. The responses were given via a conventional keyboard.

2.1.5 Analysis

Reaction times (based on within-subject median) and accuracy were entered into independent 2 (visual stimulus: circle vs. disc) x 2 (task-relevant modality: auditory vs. visual) x 2 (mapping: 1 vs. 2) factorial repeated-measures ANOVAs.

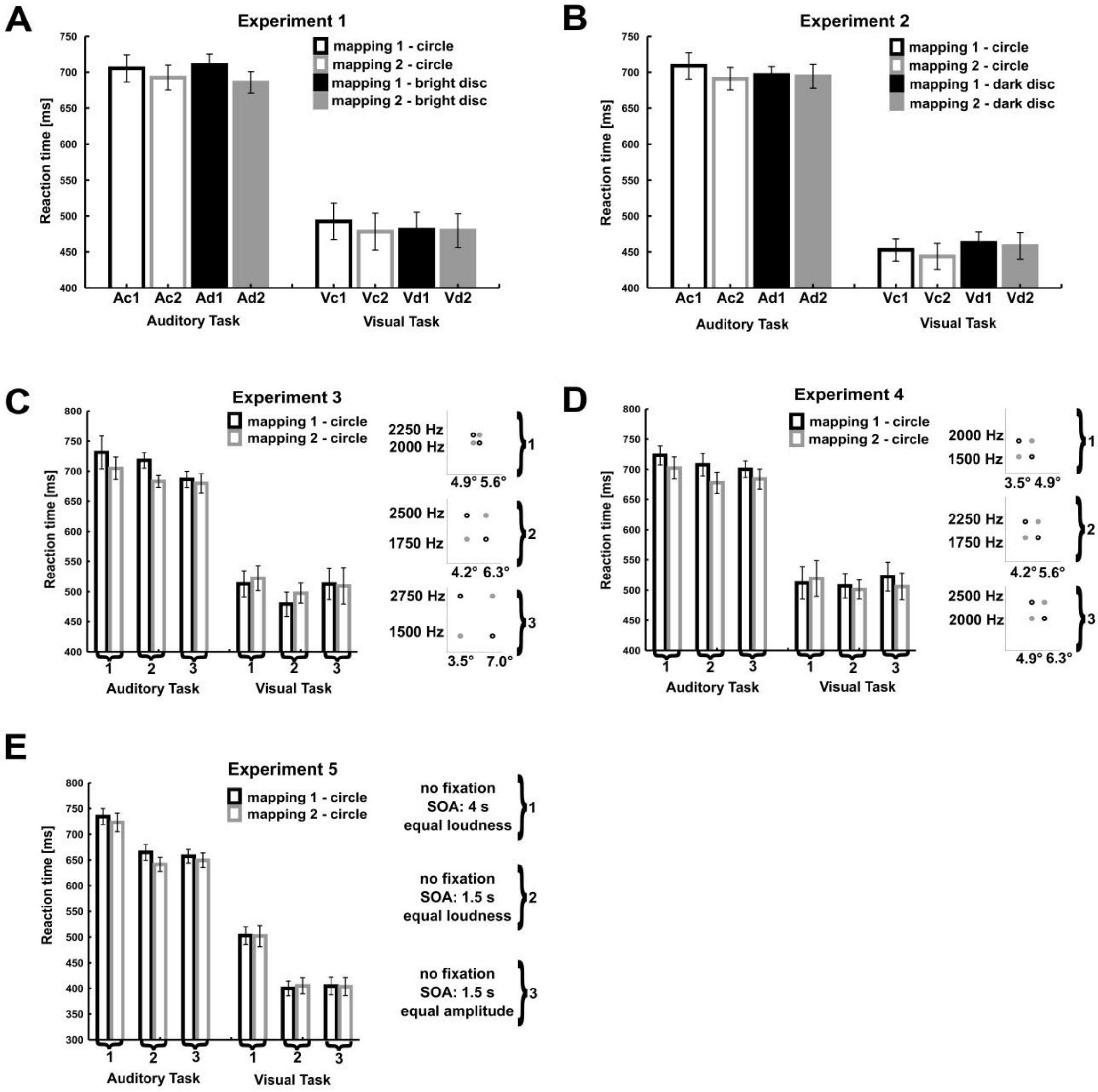


Figure 2. Experiment 1 (**A**) and 2 (**B**): Bar plots showing reaction times (across subjects' mean +/- *SEM)* for circle or disc stimuli from mapping 1 or 2. A = auditory task, V = visual task, c = circle, d = disc, 1 or 2 = mapping 1 or 2. Experiment 3 (**C**) and 4 (**D**). Bar plots showing reaction times (across subjects' mean +/- *SEM*) for circle stimuli from mapping 1 or 2. The inserts show the configuration of audio-visual combinations of the chosen two pitch and two size parameters in a space spanned by size on the x-axis and frequency on y-axis. Experiment 3 manipulated the similarity between the two stimuli: 1 = small, 2 = medium, 3 = large. Experiment 4 manipulated the relative mapping between pitch and size as shown in the inset. Experiment 5 (**E**). Experiment 5 manipulated the parameters SOA duration and sound equalization as shown in the insert.

2.2 Results

2.2.1 Experiment 1

A 2 (visual stimulus: circle vs. disc) x 2 (task-relevant modality: auditory vs. visual) and 2 (mapping: 1 vs. 2) repeated-measures ANOVA on reaction times revealed a significant main effect of mapping and task-relevant modality (Table 1). Further, it identified a significant three-way interaction between visual stimulus type, task-relevant modality and mapping.

A follow-up 2 (visual stimulus: circle or disc) x 2 (mapping: 1 vs. 2) repeated-measures ANOVA for the auditory task demonstrated that participants responded significantly faster to stimuli from mapping 2 than 1 irrespective of whether the visual stimuli were discs or circles (Table 2). For the visual task we observed a significant two-way interaction between visual stimulus and mapping (Table 2). A follow-up 2 (mapping: 1 vs. 2) repeated-measures ANOVA per visual stimulus revealed that participants were faster for mapping 2 than mapping 1 predominantly when the stimuli were circles ($F(1,15) = 8.64$, $p = .010$) but not discs ($F(1,15) = 0.04$, $p = .836$). In summary, Experiment 1 provides initial evidence that participants in our study associated low pitch with small size and high pitch with large size.

A 2 (visual stimulus: circle vs. disc) x 2 (task-relevant modality: auditory vs. visual) and 2 (mapping: 1 vs. 2) repeated-measures ANOVA on % discrimination accuracy did not reveal any significant effects (Table 1).

2.2.2 Experiment 2

A 2 (visual stimulus: circle vs. disc) x 2 (task-relevant modality: auditory vs. visual) and 2 (mapping: 1 vs. 2) repeated-measures ANOVA on reaction times revealed a significant main effect of mapping and task-relevant modality (Table 1). Participants responded faster during the visual than the auditory task. Most importantly, they responded faster to stimuli that combined low pitch tones with small circles/discs or high pitch tones with large circles/discs (mapping 2), than stimuli that combined low pitch tones with large circles/discs or high pitch tones with small circles/discs (mapping 1). Thus, Experiment 1 and 2 provide convergent evidence that participants associate low pitch with small size and high pitch with large size.

A 2 (visual stimulus: circle vs. disc) x 2 (task-relevant modality: auditory vs. visual) and 2 (mapping: 1 vs. 2) repeated-measures ANOVA on % discrimination accuracy did not reveal any significant effects, only a marginally significant effect of task-relevant modality (Table 1).

2.2.3 Comparison Experiment 1 vs. 2: Visual discs

To further investigate whether differences in overall luminance between large and small visual stimuli may contribute to the pitch-size mapping, we directly compared experiments 1 and 2 in a 2 (stimulus luminance: high vs. low) x 2 (mapping: 1 vs. 2) x 2 (task-relevant modality: auditory vs. visual) repeated-measures ANOVA limited to the conditions where discs were presented with different luminance (n.b. the circle stimuli were identical in the two Experiments). This repeated-measures ANOVA replicated the main effects of task-relevant modality ($F(1,24) = 510.32$, $p < .001$) and mapping ($F(1,24) = 4.37$, $p = .047$). In addition, we also observed a marginally significant two-way interaction between task-relevant modality and mapping ($F(1,24) = 3.13$, $p = .090$). These results raise the possibility that changes in overall luminance may interfere with size-pitch relationships.

3. Experiment 3 & 4

3.1 Introduction

Surprisingly, the results of Experiment 1 and 2 provided consistent evidence for a congruency pattern that is opposite to the profile previously described in the literature (Bien et al., 2012; Eitan et al, 2014; Evan & Treisman, 2010; Gallace & Spence, 2006; Marks et al., 1987; Mondloch & Maurer, 2004; Parise & Spence, 2009; Parise & Spence, 2012). In contrast to previous reports we observed faster responses for low pitch/small visual size and high pitch/large visual size stimuli (mapping 2). In order to test whether the opposite results might be caused by our choice of size and pitch parameters, we explored the space of parameters governing the pitch-size relationship. First, we manipulated the similarity between the two size-pitch stimuli that needed to be discriminated. We hypothesized that congruency/interference effects would be stronger when stimuli are more difficult to discriminate (i.e. the stimuli of the two classes differ less in pitch/size). Second, we varied the relative mapping between pitch and size across different runs. As a consequence, pitch-size pairings that are classified as low/small in one condition are classified as high/large in a different condition. If human observers associate pitch and size in an absolute fashion, some mappings may be more effective than others.

As Experiment 1 and 2 did not reveal any significant differences between circles and discs, we focussed on circles to limit the associated changes in overall luminance.

3.2 Methods

3.2.1 Participants

Ten participants (6 female, mean age: 24 years) with no history of neurological or psychiatric illness participated in both experiment 3 and 4 after giving written informed consent. All of them had normal hearing and normal or corrected-to-normal vision. One participant was excluded from the analysis because he/she pushed all buttons of the keyboard in a random fashion. The study was approved by the local research and ethics committee.

3.2.2 Stimuli

Visual stimuli were circles (thickness: 0.5° visual angle) with radii of 3.5°, 4.2°, 4.9°, 5.6°, 6.3° and 7°. Auditory stimuli were pure tones of 120 ms duration with frequencies of 1500 Hz, 1750 Hz, 2000 Hz, 2250 Hz, 2500 Hz and 2750 Hz, sampled at 44100 Hz, with linear onset and offset ramps of 1 ms to avoid auditory clicks.

3.2.3 Experimental design

This experimental series included two experiments with 3 runs. The experimental design in each run conformed to 2 (task-relevant modality: auditory vs. visual) by 2 (mapping: 1 vs. 2) factorial design.

Experiment 3

Experiment 3 investigated the effect of similarity between the two pitch-size stimuli that needed to be discriminated in separate runs. Specifically, we included runs with three different class similarities: (i) small with sound frequency: 2000 Hz vs. 2250 Hz and circle radius: 4.9° vs. 5.6°, (ii) medium with sound frequency: 1750 Hz vs. 2500 Hz and circle radius: 4.2° vs. 6.3°, and (iii) large with sound frequency: 1500 Hz vs. 2750 Hz and circle radius: 3.5° vs. 7°.

Experiment 4

Experiment 4 manipulated the relative mapping of high/low sound frequency and the small/large visual size in separate runs while holding the similarity between the two stimulus classes constant. Specifically, we included runs where the high/low sound frequencies and small/large circles were sampled from three different ranges: (i) sound frequency: 1500 Hz vs. 2000 Hz and circle radius: 3.5° vs. 4.9°, (ii) sound frequency: 1750 Hz vs. 2250 Hz and circle radius: 4.2° vs. 5.6°, and (iii) sound frequency: 2000 Hz vs. 2500 Hz and circle radius:

4.9° vs. 6.3°. Hence, the stimulus pairing 2000 Hz / 4.9° was classified as high in condition 1, but as low in condition 3. If participants consider size-pitch not only in relative terms, but also have some absolute scale, the results may depend on the exact pitch-size pairing.

At the beginning of the experimental session, participants were familiarized with 16 trials of the first parameter setting. Afterwards each new parameter setting was introduced by sequentially displaying the auditory and visual stimuli to ensure that the labels for 'low' and 'high' pitch and 'small' and 'large' circles were assigned correctly. Each run included 10 auditory and 10 visual task blocks. The order of the blocks was permuted. The starting modality of the first block and the order of the parameter settings were counterbalanced across participants. Unlike in Experiment 1 and 2, we reduced the number of mapping options to allow for counterbalancing with a smaller number of subjects. Subjects responded with the left hand on the visual task and with the right hand on the auditory task. The mapping of response buttons was fully counterbalanced across participants resulting in four response mapping options. Otherwise, the experimental procedures and apparatus were identical to Experiment 1 and 2.

3.2.4 Analysis

For each Experiment, reaction times (based on within-subject median) and accuracy were entered into independent n (n x parameter settings) x 2 (task-relevant modality: auditory vs. visual) x 2 (mapping: 1 vs. 2) factorial repeated-measures ANOVAs.

3.3 Results

3.3.1 Experiment 3

A 3 (similarity between stimulus classes: small, medium, large) x 2 (task-relevant modality: auditory vs. visual) and 2 (mapping: 1 vs 2) repeated-measures ANOVA on reaction times revealed a significant main effect of task-relevant modality and a significant two-way interaction between task-relevant modality and mapping (Table 4). Moreover, we observed a marginally significant main effect of class similarity indicating that stimulus discriminability influences audiovisual congruency/interference effects and a marginally significant two-way interaction between parameter options and mapping. A follow-up 3 (class similarity: small, medium, large) x 2 (mapping: 1 vs. 2) repeated-measures ANOVA for the auditory task showed that participants responded significantly faster to stimuli from mapping 2 than to stimuli from mapping 1 (Table 5). The main effect of class similarity was only marginally significant (Table 5). For the visual task these effects were only marginally significant (Table 5).

A 3 (class similarity: small, medium, large) x 2 (task-relevant modality: auditory vs. visual) and 2 (mapping: 1 vs. 2) repeated-measures ANOVA on % discrimination accuracy did not reveal any significant effects (Table 4).

3.3.2 Experiment 4

A 3 (parameter mapping: 1, 2, 3) x 2 (task-relevant modality: auditory vs. visual) and 2 (mapping: 1 vs. 2) repeated-measures ANOVA on reaction times revealed a main effect of task-relevant modality and a marginally significant effect of mapping (Table 4). These results suggest that irrespective of the exact mapping between pitch and size at least within the range tested participants were again slower to respond to stimuli from mapping 1 than mapping 2.

A 3 (parameter mapping: 1, 2, 3) x 2 (task-relevant modality: auditory vs. visual) and 2 (mapping: 1 vs. 2) repeated-measures ANOVA on % discrimination accuracy did not reveal any significant effects (Table 4).

4. Experiment 5

4.1 Introduction

In a series of control experiments we investigated whether the fixation instructions, perceptual loudness or stimulus onset asynchrony could explain the discrepancy between our results and previous reports (Bien et al., 2012; Eitan et al, 2014; Evan & Treisman, 2010; Gallace & Spence, 2006; Marks et al., 1987; Mondloch & Maurer, 2004; Parise & Spence, 2009; Parise & Spence, 2012). First, we removed the fixation cross and fixation instructions, while all other factors were identical to Experiment 1. Next, we equated the two sounds with respect to loudness (Suzuki & Takeshima, 2004). Finally, we increased the SOA to 4000 ms to prevent participants from perceiving or interpreting the change in size/pitch across successive stimuli as dynamic motion.

Like in Experiment 3 and 4 we focussed on circles to control for changes in overall luminance.

4.2 Methods

4.2.1 Participants

Sixteen participants (10 female, mean age: 26 years) with no history of neurological or psychiatric illness participated in this study after giving written informed consent. All of them had normal hearing and normal or corrected-to-normal vision. The study was approved by the local research and ethics committee.

4.2.2 Stimuli

Visual stimuli were circles (thickness: 0.5° visual angle) with radii of 2.8° and 7.7° - identical to Experiment 1 and 2 except for absence of fixation cross. Auditory stimuli were pure tones of 120 ms duration with frequencies of 1250 Hz and 3000 Hz, sampled at 44100 Hz, with linear onset and offset ramps of 1 ms to avoid auditory clicks - identical to Experiment 1 and 2 The sounds were corrected for equal loudness by presenting the 1250 Hz tone at 70 dB and the 3000 Hz tone at 65 dB (Suzuki & Takeshima, 2004).

4.2.3 Experimental design

The experimental design conformed to 3 (parameter options: long SOA and equal loudness, short SOA and equal loudness, short SOA and equal amplitude), 2 (task-relevant modality: auditory vs. visual) by 2 (mapping: 1 vs. 2) factorial design.

At the beginning of the experiment, participants were familiarized in 16 example trials. Each run included 12 auditory and 12 visual task blocks. The order of the blocks was permuted. The starting modality of the first block was counterbalanced across participants. The parameter options were presented in the following order: first, long SOA and equal loudness, second, short SOA and equal loudness, and third, short SOA and default loudness setting. This particular order was chosen to avoid carry-over effects (e.g. dynamic perception for short SOA may be transferred to long SOA). None of the experimental runs presented a fixation cross or instructed participants to fixate. The experimental procedures and apparatus were otherwise identical to Experiments 1 and 2.

4.2.4 Analysis

Reaction times (based on within-subject median) and accuracy (% correct) were entered into independent 3 (3 x parameter options: long SOA and equal loudness, short SOA and equal loudness, short SOA and equal amplitude) x 2 (task-relevant modality: auditory vs. visual) x 2 (mapping: 1 vs. 2) factorial repeated-measures ANOVAs.

4.3 Results

A 3 (parameter options: long SOA and equal loudness, short SOA and equal loudness, short SOA and equal amplitude) x 2 (task-relevant modality: auditory vs. visual) and 2 (mapping: 1 vs 2) repeated-measures ANOVA of reaction times revealed a significant main effects of parameter settings and task-relevant modality and a marginally significant main effect of mapping. Furthermore, it revealed a significant two-way interaction between task-relevant modality and mapping (Table 6). A follow-up 3 (parameter options: long SOA and equal loudness, short SOA and equal loudness, short SOA and equal amplitude) x 2 (mapping: 1 vs. 2) repeated-measures ANOVA for the auditory task showed that participants responded significantly faster to stimuli from mapping 2 than to stimuli from mapping 1 (Table 7). For the visual task these effects were not significant (Table 7). Furthermore, in both the auditory and visual task participants responded faster for short than long SOAs (Table 7

A 3 (parameter options: long SOA and equal loudness, short SOA and equal loudness, short SOA and equal amplitude) x 2 (task-relevant modality: auditory vs. visual) and 2 (mapping: 1 vs. 2) repeated-measures ANOVA of % discrimination accuracy revealed significant main effects of parameter options and mapping. Furthermore, it revealed a significant interaction between task and mapping (Table 6). In a follow-up 2 (task-relevant modality: auditory vs. visual) x 2 (mapping: 1 vs. 2) repeated measures ANOVA revealed a significant main effect of mapping indicating that participants responded more accurate to stimuli from mapping 2 than to stimuli from mapping 1 (Table 8).

5. Discussion

This study revisited the synesthetic relationship between auditory pitch and visual size. In previous research participants were faster to discriminate between different sizes in the visual modality, when small-sized stimuli were presented with high-pitched tones and large-sized stimuli with low-pitched tones. Yet, a recent study challenged this generic pitch-size mapping by demonstrating the reverse relationship for dynamic stimuli. In the dynamic context, increases in size were associated with rising pitch. To shed further light on this seemingly paradoxical finding, we have investigated several factors that can potentially influence the size-pitch mapping in static contexts.

First of all, we investigated whether stimulus luminance may have contributed to the size-pitch association. In past research luminance and size were correlated, because the overall luminance of the presentation screen will decrease for larger-grating or grey-disc stimuli when presented on a white background (Evans & Treisman, 2010; Gallace & Spence; 2006). To dissociate the effects of luminance and size, Experiments 1 and 2 compared discs that were either brighter or darker than the colour of the background. Moreover, we included circles that limit changes in overall luminance induced by changes in stimulus size. Irrespective of the stimulus (i.e. disc or circle) or the relative luminance between stimulus and background we observed faster reaction times when small size was associated with low pitch and large size with high pitch. A direct comparison between Experiment 1 and 2 raised the possibility that changes in overall luminance play a role in the pitch-size relationship. Nevertheless, changes in luminance did not revert the profile. Instead, both Experiments 1 and 2 provided convergent evidence for a pitch-size mapping that is opposite to the one previously reported in the literature (Bien et al., 2012; Eitan et al, 2014; Evan & Treisman, 2010; Gallace & Spence, 2006; Marks et al., 1987; Mondloch & Maurer, 2004; Parise & Spence, 2009; Parise & Spence, 2012).

In Experiment 3 and 4 we therefore aimed to identify additional factors that may influence how participants associate pitch and size during speeded reaction time tasks. In particular, we asked in Experiment 3 whether the pitch-size mapping depends on similarity between the two stimulus classes. We expected that the congruency effects would be stronger when the discriminability and similarity between the two stimuli is smaller. Indeed, a concurrent visual stimulus exerted a stronger influence on participants' auditory discrimination when the stimulus classes were closer together. This finding reflects the fact that multisensory influences are most pronounced and relevant when participants' perceptual and decisional uncertainty is high. For instance, participants will be more uncertain on their auditory discrimination judgment, when the two auditory signals are close in frequency space (Grinband, Hirsch & Ferrera, 2006). Critically, however, we still observed a marginally significant effect of mapping. In particular, in the auditory discrimination task participants were slower to respond to stimuli pairing low pitch with large size or high pitch with small size. Most likely, the reaction time effects were less reliably found when the visual modality is relevant, because the overall processing times were shorter. Thereby, the interfering or facilitating auditory stimulus exerted only limited impact on the visual discrimination tasks.

In Experiment 4, we finally manipulated the relative mapping between size and pitch, as human observers may potentially have an absolute AV mapping. In that case, congruency/interference effects may not only depend on the relative size and pitch of the two stimuli that need to be discriminated, but also on the absolute pitch and size pairings. However, for our parameter selection we did not observe any evidence in favour of an absolute pitch-size pairing. Replicating the results of our previous studies, we again found a significant main effect of mapping. Participants were faster to respond to stimuli pairing low pitch/small size or high pitch/large size than the opposite pairing.

In Experiment 5, we investigated the effects of SOA and perceptual loudness and fixation instruction. Even though all these additional three experiments did not instruct participants to fixate the centre of the screen, the three experiments again revealed faster response times for the small/low-pitch and large/high-pitch mapping when the auditory modality was task-relevant. Moreover, we hypothesized that stimuli with short SOA may generate a dynamic setting and thereby influence participants' preferred mapping. Yet, SOA did not influence participants' preferred mapping. Likewise, equating stimuli with respect to their perceptual loudness did not affect participant's response time profile.

In summary, all experimental series provided convergent evidence for a pitch-size mapping that pairs low pitch with small size and high pitch with large size. This is a surprising finding as it contradicts previous findings in the literature. Moreover, it is inconsistent with the natural association between the size of a resonance body and the frequency of the sounds it produces.

We suggest that the key for understanding these seemingly contradictory results lies in how participants interpret stimulus size. Crucially, retinotopic size is determined by two key factors. First, it depends on the constant size of the stimulus in the environment. Second, it depends on the distance of the stimulus from the observer. The stimuli in our study – in particular the circles - were less likely to be associated with different objects, but rather with one object at different distances from the observer. Participants may also have performed the task by comparing the current stimulus implicitly to previous ones and judging whether it was closer or farther away. In this way, our study links the previous findings on the pitch-size mapping under static and dynamic contexts. If size is interpreted as the size of an object or a resonance body, large size is associated with low pitch as previously reported in the literature for static contexts. However, if size is interpreted as distance from the observer as in the current study and previous dynamic contexts (Eitan et al., 2014), large size is associated with high pitch. Future studies are needed to carefully manipulate participants' interpretation of 'size' as object size or distance from observer. For instance, experiments may manipulate instructions, background story or change the stimuli to guide participants' interpretation towards either object size or distance from the observer.

Moreover, even though the absolute pitch and size values did not significantly affect participants' response time profile in experiment 4, this finding may not generalize to the entire range of pitch frequency and size values. For instance, it is conceivable that small circles map to high pitch and large circles to low pitch outside the tested range of values.

Anecdotally, some of our participants mentioned that they perceived both sounds as high-pitched in our experiments. In other words, even though participants do not have absolute pitch, they may still be endowed with some coarse pitch classification scheme. If both the pitch-levels chosen violate participants' coarse pitch classification, audiovisual interference experiments may be attenuated or even reverted.

Finally, future research also needs to further investigate the role of sound amplitude and perceived loudness. In our experiments we equated sounds of different frequencies with respect to their physical sound amplitude (experiment 1-4) or perceived loudness (experiment 5, run 1-2) based on published equal loudness contours. However, equal loudness contours may differ between subjects. Hence, future studies are required that carefully equate sound loudness individually for each participant (e.g. using adaptive staircases).

In conclusion, this series of AV interference experiments showed that participants map small size onto low pitch and large size onto high pitch under specific parameter settings when the auditory modality was task relevant. These results suggest that the pitch-size mapping may be less generic and stable than previously assumed. It may depend on the exact stimulus parameters, task-context and potentially prior experience of the participant.

6. References

Adam, R., & Noppeney, U. (2010). Prior auditory information shapes visual category-selectivity in ventral occipito-temporal cortex. *NeuroImage*, *52*(4), 1592–1602.

Ashley, R. (2004). Musical pitch space across modalities: spatial and other mappings through language and culture. *Proceedings of the 8th International Conference on Music Perception & Cognition*, 64–71.

Antovic, M. (2009). Musical Metaphors in Serbian and Romani Children: An Empirical Study. *Metaphor and Symbol*, *24*(3), 184–202.

Ben-Artzi, E., & Marks, L. E. (1995). Visual-auditory interaction in speeded classification: role of stimulus difference. *Perception & Psychophysics*, *57*(8), 1151–1162.

Bernstein, I. H., & Edelstein, B. A. (1971). Effects of some variations in auditory input upon visual choice reaction time. *Journal of Experimental Psychology*, *87,* 241–247.

Bien, N., ten Oever, S., Goebel, R., & Sack, A. T. (2012). The sound of size: crossmodal binding in pitch-size synesthesia: a combined TMS, EEG and psychophysics study. *NeuroImage*, *59*(1), 663–72.

Brang, D., Rouw, R., Ramachandran, V. S., & Coulson, S. (2011). Similarly shaped letters evoke similar colors in grapheme-color synesthesia. *Neuropsychologia*, *49*(5), 1355–1358.

Donohue, S. E., Roberts, K. C., Grent-'t-Jong, T., & Woldorff, M. G. (2011). The cross-modal spread of attention reveals differential constraints for the temporal and spatial linking of visual and auditory stimulus events. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *31*(22), 7982–7990.

Eitan, Z., Schupak, A., Gotler, A., & Marks, L. E. (2014). Lower pitch is larger, yet falling pitches shrink. *Experimental Psychology*, *61*(4), 273–284.

Evans, K. K., & Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features, *Journal of Vision, 10*, 1–12.

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.

Gallace, A., & Spence, C. (2006). Multisensory synesthetic interactions in the speeded classification of visual size. *Perception & Psychophsics*, *68*(7), 1191–1203.

Grinband, J., Hirsch, J., & Ferrera, V. P. (2006). A neural representation of categorization uncertainty in the human brain. *Neuron*, *49*(5), 757–763.

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 and Development*, *35*(4), 727–732.

Kim, K.-H., & Iwamiya, S.-I. (2014). Formal Congruency between Telop Patterns and Sound Effects. *Music Perception*, *25*(5), 429-448.

Laurienti, P. J., Kraft, R. a., Maldjian, J. a., Burdette, J. H., & Wallace, M. T. (2004). Semantic congruence is a critical factor in multisensory behavioral performance. *Experimental Brain Research*, *158*, 405–414.

Lee, H., & Noppeney, U. (2011). PNAS Plus: Long-term music training tunes how the brain temporally binds signals from multiple senses. *Proceedings of the National Academy of Sciences*, *108*(51), 1441–1450.

Lee, H., & Noppeney, U. (2014). Music expertise shapes audiovisual temporal integration windows for speech, sinewave speech, and music. *Frontiers in Psychology*, *5*(August), 1–9.

Lewis, R., & Noppeney, U. (2010). Audiovisual synchrony improves motion discrimination via enhanced connectivity between early visual and auditory areas. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(37), 12329–12339.

Macaluso, E., & Driver, J. (2005). Multisensory spatial interactions: A window onto functional integration in the human brain. *Trends in Neurosciences*, *28*(5), 264–271.

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., Hammeal, R. J., Bornstein, M. H., & Smith, L. B. (1987). Perceiving similarity and comprehending metaphor. *Monographs of the Society for Research in Child Development*, *52*(1, Whole No. 215), 1-102.

Marks, L. E. (2004). Cross-modal interactions in speeded classification. In G. A. Calvert, C. Spence, & B. E. Stein (Eds.), *Handbook of multisensory processes* (pp. 85–105). Cambridge: MIT Press.

Martino, G., & Marks, L. E. (1999). Perceptual and linguistic interactions in speeded classification: Tests of the semantic coding hypothesis. *Perception*, *28*(7), 903–923.

Melara, R.D., & O'Brien, T.P. (1987). Interaction between synesthetically corresponding dimensions. *Journal of Experimental Psychology: General, 116,* 323–336.

Mondloch, C. J., & Maurer, D. (2004). Do small white balls squeak? Pitch-object correspondences in young children. *Cognitive, Affective & Behavioral Neuroscience*, *4*(2), 133–6.

Parise, C. V., & Spence, C. (2009). "When birds of a feather flock together": synesthetic correspondences modulate audiovisual integration in non-synesthetes. *PloS One*, *4*(5), e5664.

Parise, C. V, & Spence, C. (2012). Audiovisual crossmodal correspondences and sound symbolism: a study using the implicit association test. *Experimental Brain Research*, *220*(3-4), 319–333.

Parise, C. V, Knorre, K., & Ernst, M. O. (2014). Natural auditory scene statistics shapes human spatial hearing. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 6104–6108.

Patching, G. R., & Quinlan, P. T. (2002). Garner and congruence effects in the speeded classification of bimodal signals. *Journal of Experimental Psychology. Human Perception and Performance*, *28*(4), 755–775.

Rich, a. N., Bradshaw, J. L., & Mattingley, J. B. (2005). A systematic, large-scale study of synaesthesia: Implications for the role of early experience in lexical-colour associations. *Cognition*, *98*(1), 53–84.

Sadaghiani, S., Maier, J. X., & Noppeney, U. (2009). Natural, metaphoric, and linguistic auditory direction signals have distinct influences on visual motion processing. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*(20), 6490–6499.

Spence, C. (2011). Crossmodal correspondences: a tutorial review. *Attention, Perception & Psychophysics*, *73*(4), 971–995.

Spence, C., & Deroy, O. (2013). How automatic are crossmodal correspondences? *Consciousness and Cognition*, *22*(1), 245–260.

Suzuki, Y., & Takeshima, H. (2004). Equal-loudness-level contours for pure tones. *The Journal of the Acoustical Society of America*, *116*(2), 918–933.

Van Atteveldt, N., Formisano, E., Goebel, R., & Blomert, L. (2004). Integration of letters and speech sounds in the human brain. *Neuron*, *43*, 271–282.

Van Atteveldt, N. M., Formisano, E., Blomert, L., & Goebel, R. (2007). The effect of temporal asynchrony on the multisensory integration of letters and speech sounds. *Cerebral Cortex*, *17*(April), 962–974.

Von Kriegstein, K., Smith, D. R. R., Patterson, R. D., Ives, D. T., & Griffiths, T. D. (2007). Neural representation of auditory size in the human voice and in sounds from other resonant sources. *Current Biology : CB*, *17*(13), 1123–1128.

Vroomen, J., & Keetels, M. (2010). Perception of intersensory synchrony: A tutorial review. *Attention, Perception & Psychophsics*, *72*(4), 871-884.

Wallace, M. T. Wilkinson, L. K., & Stein, B. E. (1996). Representation and integration of multiple sensory inputs in primate superior colliculus. *Journal of Neurophysiology, 76(2),* 1246-1266.

Wallace, M. T., Roberson, G. E., Hairston, W. D., Stein, B. E., Vaughan, J. W., & Schirillo, J. (2004). Unifying multisensory signals across time and space. *Experimental Brain Research*, *158*, 252–258.

Table 1. Statistical results of Experiment 1 and 2

	Experiment 1 (df: 1,15)		Experiment 2 (df: 1,9)	
	reaction time	accuracy	reaction time	accuracy
visual stimulus	F = 0.25	F = 0.00	F = 0.36	F = 0.38
	p = .623	p = .993	p = .566	p = .552
task-relevant modality	F = 255.79	F = 2.12	F = 1444.61	F = 4.77
	p < .001*	p = .166	p < .001*	p = .057
mapping	F = 11.83	F = 0.21	F = 5.75	F = 0.06
	p = .004*	p = .653	p = .040*	p = .817
visual stimulus x task-relevant modality	F = 0.18	F = 0.52	F = 0.90	F = 2.67
	p = .678	p = .480	p = .367	p = .137
visual stimulus x mapping	F = 0.07	F = 0.03	F = 0.76	F = 0.55
	p = .803	p = .867	p = .407	p = .476
task-relevant modality x mapping	F = 4.42	F = 0.08	F = 0.23	F = 0.56
	p = .053	p = .785	p = .640	p = .473
visual stimulus x task-relevant modality x mapping	F = 7.34	F = 0.59	F = 0.26	F = 0.25
	p = .016*	p = .453	p = .625	p = .628

**p* < 0.05

Table 2. Statistical results of follow-up 2-way ANOVAs on reaction time for auditory and visual task of Experiment 1

	Experiment 1 (df: 1,15)	
	auditory task	visual task
visual stimulus	F = 0.02	F = 0.37
	p = .884	p = .552
mapping	F = 15.22	F = 3.21
	p < .001*	p = .093
visual stimulus x mapping	F = 3.29	F = 4.57
	p = .090	p = .049*

**p* < 0.05

Table 4. Statistical results of Experiment 3 and 4

	Experiment 3		Experiment 4	
	reaction time	accuracy	reaction time	accuracy
parameter options	F(1.68,13.45) = 3.69 p = .059	F(1.82,14.56) = 0.09 p = .904	F(1.56,12.37) = 1.07 p = .356	F(1.29,10.31) = 0.06 p = .866
task-relevant modality	F(1,8) = 144.16 p < .001*	F(1,8) = 1.02 p = .325	F(1,8) = 132.93 p < .001*	F(1,8) = 2.40 p = .160
mapping	F(1,8) = 4.65 p = .063	F(1,8) = 2.23 p = .173	F(1,8) = 5.33 p = .050*	F(1,8) = 0.17 p = .694
parameter options x task-relevant modality	F(1.96,15.68) = 3.61 p = .052	F(1.61,12.91) = 0.78 p = .454	F(1.77,14.19) = 0.70 p = .496	F(1.56,12.51) = 0.19 p =.778
parameter options x mapping	F(1.74,13.93) = 0.07 p = .912	F(1.96,15.67) = 0.32 p = .724	F(1.85,14.78) = 0.51 p = .596	F(1.30,10.39) = 1.05 p =.352
task-relevant modality x mapping	F(1,8) = 15.11 p = .005*	F(1,8) = 1.12 p = .321	F(1,8) = 2.29 p = .169	F(1,8) = 1.79 p =.218
parameter options x task-relevant modality x mapping	F(1.40,11,18) = 2.78 p = .116	F(1.48,11.83) = 1.89 p = .196	F(1.70,13.59) = 1.63 p = .226	F(1.28,10.17) = 1.05 p =.350

**p* < 0.05

Table 5. Statistical results of follow-up 2-way ANOVAs on reaction time for auditory and visual task of Experiment 3

	Experiment 3	
	auditory task	visual task
similarity	*F*(1.05,8.39) = 4.15 *p* = .073	*F*(1.70,13,62) = 3.17 *p* = .080
mapping	*F*(1,8) = 14.12 *p* = .006*	*F*(1,8) = 3.79 *p* = .087
similarity x mapping	*F*(2.00,16.00) = 1.80 *p* = .196	*F*(1.59,12.69) = 0.97 *p* = .387

**p* < 0.05

Table 6. Statistical results of Experiment 5

	Experiment 5	
	reaction time	accuracy
parameter options	F(1.29,19.33) = 70.61 p < .001*	F(1,15) = 21.43 p <.001*
task-relevant modality	F(1,15) = 912.66 p <.001*	F(1,15) = 2.81 p = .115
mapping	F(1,15) = 4.45 p = .052	F(1,15) = 5.57 p = .032*
parameter options x task-relevant modality	F(1.61,24.17) = 3.20 p = .068	F(1,15) = 2.81 p = .115
parameter options x mapping	F(1.70,25.50) = 0.56 p = .563	F(1,15) = 5.57 p = .032*
task-relevant modality x mapping	F(1,15) = 4.86 p = .044*	F(1,15) = 0.07 p = .790
parameter options x task-relevant modality x mapping	F(1.78,26.67) = 1.88 p = .175	F(1,15) = 0.07 p = .790

*$p < 0.05$

Table 7. Statistical results of follow-up 2-way ANOVAs on reaction time for auditory and visual task of Experiment 5

	Reaction Time	
	auditory task	visual task
parameter options	F(1.37,20.47) = 35.6 p < .001*	F(1.39,20.83) = 68.83 p < .001*
mapping	F(1,15) = 7.12 p = .018*	F(1,15) = 0.07 p = .794
parameter options x mapping	F(1.58,23.68) = 2.57 p = .107	F(1.71,25.78) = 0.402 p = .672

*$p < 0.05$

Table 8. Statistical results of follow-up 2-way ANOVAs on accuracy per parameter option of Experiment 5

	Accuracy	
	short SOA, equal amplitude	short and long SOA, equal loudness
task-relevant modality	F(1,15) = 2.81 p = .115	No effect: 100% correct
mapping	F(1,15) = 5.57 p = .032*	
task-relevant modality x mapping	F(1,15) = 0.07 p = .790	

*$p < 0.05$

Highlights:

- Intersensory selective attention paradigm with auditory pitch and visual size
- Convergent evidence for small size/low pitch, large size/high pitch mapping
- Mapping concordant with previous findings in dynamic but not static context
- Interpretation of changing size possible as object size or distance from observer
- Direction of pitch-size mapping depends on parameter choice and task instructions