The relation of space and musical pitch in the brain

Sarah Dolscheid [1,2,3]
(sarah_dolscheid@gmx.de)

Roel M. Willems [1,4]
(roel.willems@fcdonders.ru.nl)

Peter Hagoort [1,4]
(Peter.Hagoort@mpi.nl)

Daniel Casasanto [5]
(casasanto@uchicago.edu)

[1]Max Planck Institute for Psycholinguistics, Nijmegen, NL
[2]International Max Planck Research School for Language Sciences, Nijmegen, NL
[3]Department of Rehabilitation and Special Education, University of Cologne, GER
[4]Donders Institute for Brain, Cognition, and Behaviour, Radboud University, Nijmegen, NL
[5]Department of Psychology, University of Chicago, USA

Abstract

Numerous experiments show that space and musical pitch are closely linked in people's minds. However, the exact nature of space-pitch associations and their neuronal underpinnings are not well understood. In an fMRI experiment we investigated different types of spatial representations that may underlie musical pitch. Participants judged stimuli that varied in spatial height in both the visual and tactile modalities, as well as auditory stimuli that varied in pitch height. In order to distinguish between unimodal and multimodal spatial bases of musical pitch, we examined whether pitch activations were present in modality-specific (visual or tactile) versus multimodal (visual and tactile) regions active during spatial height processing. Judgments of musical pitch were found to activate unimodal visual areas, suggesting that space-pitch associations may involve modality-specific spatial representations, supporting a key assumption of embodied theories of metaphorical mental representation.

Keywords: pitch, space, metaphor, embodiment, fMRI

Introduction

According to theories of metaphorical mental representation, many of our abstract concepts are scaffolded by spatial schemas (Lakoff & Johnson, 1980). In behavioral experiments, spatial representations appear to contribute to people's understanding of domains like time (Boroditsky, 2000), emotional valence (Meier & Robinson, 2004), power (Schubert, 2005), and similarity (Casasanto, 2008): domains that can never be perceived with the senses. Spatial representations also seem to underlie some domains that can be perceived directly, like musical pitch.

Close links between spatial height and pitch "height" have been demonstrated in various ways. In speeded classification tasks, for instance, people respond faster to a stimulus when visuospatial height and pitch are congruent than when they were incongruent (e.g., Melara & O'Brien, 1987; Evans & Treisman, 2010). Participants also press response keys that are spatially high more quickly in response to high-frequency pitches than in response to low-frequency pitches (and vice versa for spatially low response keys), as shown in stimulus-response compatibility experiments (Lidji, Kolinsky, Lochy, & Morais, 2007; Rusconi, Kwan, Giordano, Umiltà, & Butterworth, 2006). Even prelinguistic infants have been found to be sensitive to height-pitch associations (e.g., Walker et al., 2010;

Dolscheid, Hunnius, Casasanto, & Majid, 2012; Jeschonek, Pauen, & Babocsai, 2012). Although numerous behavioral experiments confirm links between spatial height and musical pitch, they do not shed light on the neuronal underpinnings of space-pitch associations.

According to theories of embodied cognition, neural systems for perception and action also subserve thinking. When people perceive stimuli, representations in modality-specific brain areas (e.g., visual cortex, auditory cortex) are captured by conjunctive neurons in multimodal cortical association areas, forming *simulators* (Barsalou, Simmons, Barbey & Wilson, 2003). When information captured by these simulators is needed later, these multimodal conjunctive neurons activate modality-specific neuronal populations, partially recreating perceptual states in the absence of any sensory input: a process called *simulation*.

Whereas modality-specific simulation has been demonstrated for a number of cognitive domains (e.g., motion, Saygin, McCullough, Alac, & Emmorey, 2010), little is known about simulation in metaphorical mental representations. Are spatial representations that underlie domains like musical pitch modality-specific? More precisely, does musical pitch rely, at least in part, on visuospatial representations, as has been suggested by some researchers (e.g., Eitan, Ornoy, & Granot, 2012)?

Some hints at an answer come from neuroscientific investigations of pitch-related tasks. In several studies, pitch processing was accompanied by activations in primary visual areas such as the cuneus and the calcarine cortex (e.g., Foster & Zatorre, 2010; Perry et al., 1999; Zatorre, Meyer, Gjedde, & Evans, 1996). Platel et al. (1997) for instance found that unexpectedly the left cuneus (in the occipital lobe) was one of the main areas active during detection of pitch changes in a sequence of sounds (Platel et al., 1997). Also Degerman and colleagues reported pitch-activated brain regions including the right cuneus (Degerman, Rinne, Salmi, Salonen, & Alho, 2006). Zatorre et al. (1996) even found primary visual activations in pitch-related tasks when participants' eyes were closed, and Perry et al. (1999) reported activity in the calcarine cortex (BA 17) during singing. Taken together these findings suggest that pitch processing might depend on some basic visual regions of the brain. Representations that underlie musical pitch may therefore indeed be partly visuospatial in nature.

Crucially, however, space is not restricted to visual perception. Rather, spatial experiences are frequently multimodal and often comprise the integration of visual, vestibular, auditory and even somatosensory cues. Multisensory regions like the inferior parietal lobe (IPL) or the intraparietal sulcus (IPS) have been found to be relevant for coding spatial experiences across a variety of modalities (e.g., Amorapanth, Widick, & Chatterjee, 2010). Moreover, Schwenzer and Mathiak (2011) reported IPS activations in a musical context, as the result of a pitch identification task. Zatorre, Halpern, and Bouffard (2010) further demonstrated IPS involvement in melody-related judgments. Participants were asked to imagine what a reversed melody line would sound like. This melodic reversal led to activation in anterior portions of the IPS (see also Foster, Halpern, & Zatorre, 2013). Similar activation was found when participants made judgments about transposed melodies (Foster & Zatorre, 2010). According to these findings, musical pitch may involve schematic representations of space, instantiated in areas of cortex that subserve multimodal or amodal spatial processing.

Are spatial representations of pitch instantiated in unimodal (e.g., visuospatial) brain areas, multimodal areas, or both? Here we examined the spatial basis of pitch in an fMRI experiment in which we directly compared the neural correlates of spatial height and pitch height. To determine whether pitch representations overlap with unimodal or multimodal spatial representations (or both), we asked participants to judge spatial height in two modalities; vision and touch. Participants were asked to decide whether two serially presented stimuli (simple shapes like circles and squares) differed in vertical position. In the visual condition, participants saw the stimuli on the screen, and in the tactile condition, participants felt the stimuli on their palms. In addition, participants were asked to judge whether two successive sounds differed in pitch. We reasoned that if pitch judgments rely on the same neural circuitry as judgments concerning visual, tactile, or multimodal space (operationalized as the intersection of visual and tactile space), we should find pitch-related BOLD signal changes in region(s) that are involved in spatial height processing.

To rule out the possibility that the overlap between regions involved in space and pitch judgments was simply due to similarity in task demands (i.e. judgment processes), control conditions were added for all three tasks. In the control conditions, participants were also asked to judge whether two serially presented stimuli differed or not. However, whereas target stimuli varied in position (or pitch), control stimuli remained constant (i.e. at the same location/pitch). To ensure that participants' responses were not biased (i.e. that the correct answer would not always be "different" during the experimental trials and always be "same" in the control trials), participants were either asked to judge whether two successive stimuli were of the same height or whether the stimuli were the same shape for both experimental and control conditions. The same strategy was applied to auditory stimuli: Participants were asked to either focus on whether two successive tones had the same musical pitch or the same timbre (i.e., played by same or different instruments).

If pitch representations are instantiated in modality-specific visual cortices, pitch judgments should activate areas that are also selectively activated by visuospatial height processing. Alternatively, if pitch representations draw on modality-specific cortices that are not restricted to vision, pitch judgments should (also) activate areas selective for tactile height processing. Finally, if pitch representations are scaffolded by spatial representations in multimodal or amodal brain areas, then pitch judgments should activate regions that are activated by both visual and tactile spatial height judgments.

On the basis of these predictions, we also defined several regions of interest (ROIs). We looked for potential pitch activations in both unimodal and multimodal ROIs. Primary visual cortex (BA 17) served as a visual ROI (Bosking, Crowley, & Fitzpatrick, 2002; Noesselt et al., 2002). The postcentral gyrus, a region involved in tactile processing (Macaluso & Driver, 2001; 2005), was selected as a tactile ROI. Finally, the IPL served as a multimodal ROI (Macaluso & Driver, 2005).

Methods

Participants

We tested 20 healthy right-handed Dutch speakers (14 women; mean age = 22.8 years, range = 18 − 30 years, 6 men; mean age = 25.7 years, range = 19 − 53 years) with no known history of neurological problems, dyslexia or other language-related problems or hearing complaints, and with normal or corrected-to-normal vision. All participants provided written informed consent and were compensated for their participation (10€/hour). One participant's data could not be saved due to storage problems. Three participants had to be excluded from further analyses due to large head movement during the scanning session. In total, 16 participants remained in the sample. The study was approved by the local ethical committee for research with human participants.

Materials

Three different types of materials were used in three blocks: visual, tactile and auditory.

Visual materials: Two pictures of simple white objects were presented on a black background. Stimuli consisted of a 3 cm wide circle and a 3 cm wide square (visual angle: 2.86°), either presented at the upper part of the screen (approximately 8 cm from mid of screen, visual angle: 7.63°) or at the lower part of the screen (approximately 8 cm from mid of screen, visual angle: 7.63°).

Tactile materials: Stimuli consisted of a 3 cm wide wooden circle and a 3 cm wide wooden square. Shapes were constructed such that a ridge around the perimeter of the shapes (approximately 2 mm wide) could be pressed against the participant's palms, either at a position high in tactile

space (upper part close to the fingers of the participant) or at a lower position (lower part close to the participant's wrist, as the participants flexed their right hand with the fingers pointing upward, see Figure 1).

Auditory materials: Stimuli consisted of 4 sounds. Timbres of a trumpet and a cello were produced by a Korg Triton synthesizer and were afterwards modified in Adobe Audition 1.5 (Adobe Systems Inc.). Both timbres were presented at two frequencies, to produce a low-pitched sound (262 hz) and a relatively high-pitched sound (394 hz).
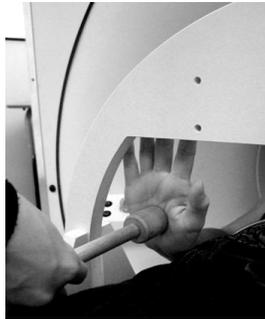
Figure 1: Tactile stimulation procedure: The participant's palm faced the experimenter, with fingers pointing upward. The wooden stimulus was pressed against the palm, either at a position high in tactile space (upper part close to the fingers of the participant) or at a lower position (lower part close to the participant's wrist).

Procedure

Visual and auditory stimuli were presented using Presentation software (www.neurobs.com, version 14.2). Instructions and visual stimuli were presented through a projector from outside the scanner room onto a screen at the back of the scanner bore and were visible to the participants through a mirror attached to the head coil. Materials were presented in 3 different blocks (visual, tactile, auditory). The order of blocks was counterbalanced across participants. Before the fMRI session, participants were familiarized with the task outside of the scanner. Participants were presented with 10 trials of each of the three blocks to illustrate the procedure.

In the visual block, at the beginning of each trial participants were presented either the word 'positie' (position) or the word 'vorm' (shape) printed on the screen (see Figure 2). Words served as a prompt to indicate the stimulus attribute that was relevant and should be attended to on a given trial (position versus shape). Afterwards, participants were asked to compare 2 visual stimuli (each lasting for 1 second on the screen) and indicate as accurately and fast as possible whether both stimuli were the same or different with respect to the relevant dimension (e.g., position). Participants responded with button presses of their left index or middle finger while response options 'hetzelfde' (same) and 'verschillend' (different) were printed on the screen. Response side was counterbalanced across participants. Overall, two types of shape (circle and square) were fully crossed with two dimensions of position (high and low). For both experimental and control trials, participants were asked to judge whether the subsequent stimuli were of the same height in half of the trials and whether they were of the same shape in the other half. In total, the visual block consisted of 64 trials, 32 experimental trials (in which position varied) and 32 control trials (in which position remained constant). The order of stimuli and trial type was randomized.

The procedure of the tactile block was identical to the visual block. At the beginning of each trial participants were presented either the word 'positie' (position) or the word 'vorm' (shape) printed on the screen. However, this time the modality of the stimuli was tactile. Tactile stimuli were operated by the experimenter who was wearing headphones. Presentation of different beeps to the headphones indicated the timing and dimension of the tactile stimulus (high vs. low, circle vs. square). Participants were asked to stretch their right arm in parallel to their body. The right hand was supported by a wooden arch placed over the participant's abdomen, with the palm facing the experimenter and fingers pointing towards the ceiling (see Figure 1). Prompted by the beeps (only audible to the experimenter), the experimenter touched the participant's palm with the respective tactile stimulus for around 1 second (high=close to the participant's fingers, low=close to the participant's wrist). In total, the tactile block consisted of 64 trials, 32 experimental trials and 32 control trials.

The procedure of the auditory block was identical to the visual and the tactile block. However, during the auditory block participants were presented either the word 'toon' (tone) or the word 'instrument' printed on the screen at the beginning of each trial. Words served as a prompt to indicate the stimulus attribute that was relevant on a given trial (tone=pitch vs. instrument=timbre). Participants were then asked to compare 2 subsequent auditory stimuli that were presented via scanner-compatible headphones (each lasting 1 second). Participants also wore headphones during the entire experiment to protect their hearing from scanner noise. In total, the auditory block consisted of 64 trials, 32 experimental trials and 32 control trials.
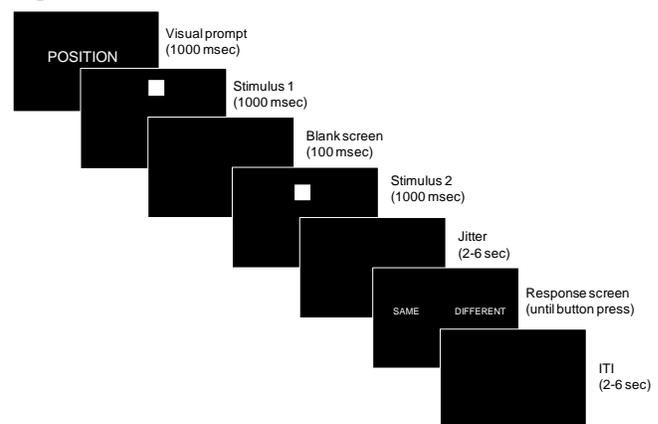
Figure 2: Example of a visual trial. Participants saw a visual prompt at the beginning of each trial. Then two stimuli were presented. Participants were asked to compare these stimuli with respect to the relevant dimension (in this example; position). After a jittered delay, participants responded with their left hand indicating whether the presented stimuli differed or not.

FMRI data analysis

Functional data were preprocessed and analyzed with SPM8 (Statistical Parametric Mapping, Wellcome Department of Cognitive Neurology, London, UK). The first 5 volumes of each functional sequence were removed to allow for T1 equilibration effects.

To correct for head movements, images were spatially realigned with rigid body registration along three translational and three rotational axes. Images were temporally realigned to correct for slice timing acquisition delays to the onset of the first slice. Next, images were coregistered to each subject's structural scan and normalized to a standard EPI template in Montreal Neurological Institute space and resampled at an isotropic voxel size of 2 mm. The normalized images were then smoothed with an isotropic 8 mm FWHM Gaussian kernel.

These preprocessed data were analyzed on a subject by subject basis using an event-related approach. The time series of the preprocessed data were entered into a General Linear Model, and separate regressors were calculated for trials in which an actual change in position/pitch took place (height change) as compared to trials in which no change occurred (control). This resulted in the following regressors: visual height change, visual control, tactile height change, tactile control, pitch change, pitch control. Only trials with a correct response were considered. Events were timed at the occurrence of the second stimulus in a trial, and were modeled as stick functions and then convolved with a canonical hemodynamic response function (HRF). Responses (button presses) were modeled separately as stick functions. Finally, the estimates of the motion correction algorithm were added as nuisance regressors to the model to account for disturbances caused by small head movement. In order to localize activity related to spatial height change in all three modalities, we computed contrast images of height change and the control condition (no change) for each participant.

A second-level whole-brain group analysis with subjects as a random factor ('random effects analysis') was carried out. Here, we looked for regions that were selectively activated by (height change-control) in vision and touch, as well as multimodal regions activated by both visual and tactile height change [visual (height change-control) ∩ tactile (height change-control)]. To correct for the number of comparisons in this massive univariate approach (multiple comparisons problem), we combined a *p*<0.001 (uncorrected) voxel threshold with a cluster extent threshold, to arrive at a corrected p-value of *p*<0.05. The cluster extent threshold was determined by reference to the theory of Gaussian Random Fields (Friston et al., 1996; Poline et al., 1997).

Given our a priori hypothesis about pitch activations in regions involved in spatial height perception, we performed two region of interest analyses. In the first ROI analysis we looked for activations of (pitch change) and (pitch change-control) in regions that were selectively activated by visual height change, tactile height change and multimodal height change. Only voxels with a *p* < .001 (whole brain, uncorrected) were considered.

In a second ROI analysis we looked at visual as well as tactile (height change-control) activations in predefined anatomical regions of interest. In case of significant visual or tactile (height change-control) activations, we also looked for activations of (pitch change) and (pitch change-control) in the activated region(s). ROIs were selected by using the WFU pickatlas (Lancaster et al., 2000; Tzourio-Mazoyer, 2002) and MarsBaR (http:// marsbar.sourceforge.net/, version 0.42). The postcentral gyrus, a unimodal region that is involved in tactile processing (Macaluso & Driver, 2001; 2005), was selected as a tactile ROI (we restricted this ROI to the left hemisphere, since tactile stimulation was only administered to participants' right hands). Primary visual cortex (BA 17) was selected as a visual ROI (Noesselt et al., 2002). Furthermore, based on previous research, the IPL was selected as multimodal ROI (e.g., Macaluso & Driver, 2005). Again, only the left hemisphere was considered since tactile stimulation was restricted to the right hand.

Results

Behavioral results

Overall, participants completed the tasks with high accuracy (Mean accuracy = 92%). In the visual task, performance was high for both the experimental trials (Mean accuracy [visual height change] = 96%) and the control trials (Mean accuracy [visual control] = 94%). In the tactile task, performance was slightly lower for both the experimental trials (Mean accuracy [tactile height change] = 86%) and the control conditions (Mean accuracy [tactile control] = 88%). In the auditory task, performance was high for both the experimental trials (Mean accuracy [pitch change] = 92%) and the control trials (Mean accuracy [pitch control] = 94%). In all three modalities, there were no significant differences in error rates between experimental and control conditions (all *p*-values *ns*).

Whole brain analyses and functional ROI analyses

Visual activations: Visual height judgments (height change-control) corresponded to significant activity in the occipital cortex [MNI coordinates: 4, -80, 18]. We used this region as a region of interest (ROI) and found that there was also significant activation of (pitch change), $t(15)=2.78$, $p=.01$. However, there were no significant activations for (pitch change-control), $t(15)=1.41$, *ns*.

Tactile activations: There were no significant activations of tactile (height change-control).

Multimodal activations: There were no significant activations of the conjunction of [visual (height change-control) ∩ tactile (height change-control)].

Anatomical ROI analyses

Visual: A ROI analysis in BA 17 (anatomically defined) revealed a significant cluster of visual (height change-control) activity [MNI coordinates: 2, -84, 10]. Within this region, there was significant activation of pitch change

t(15)=3.4, *p*=.004, whereas the auditory control condition revealed no significant activity, *t*(15)=1.6, *ns*; difference of activation, *t*(15)=2.17, *p*=.05 (Figure 3). Moreover, this cluster revealed a significant deactivation in tactile height change *t*(15)=-5.26, *p*=.0001, but no significant signal change for tactile (height change-control), *t*(15)=.18, *ns*.

Tactile: There was no significant activation of tactile (height change-control) in the postcentral ROI.

Multimodal: There was no significant activation of the conjunction [visual (height change-control) ∩ tactile (height change-control)] in the IPL.
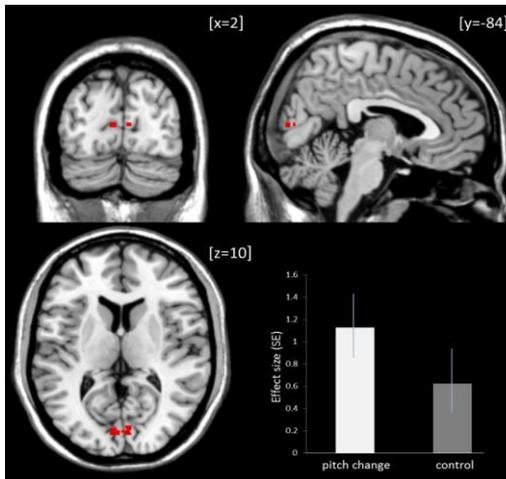
Figure 3: Significant visual (height change-control) activity in BA 17. In addition to visual height sensitivity, there was also significant activity correlated with pitch change, and with (pitch change-control).

Discussion

By comparing stimuli that differed in spatial height to those that remained at a constant position (control), we found activations in primary visual cortex. In this primary visual area we also observed activity correlated with (pitch change), however there was no significant activation for (pitch change-control) in this ROI analysis. In the anatomical ROI analysis, however, a cluster restricted to BA17 specifically responding to visual (height change-control) also revealed significant activity for (pitch change) as well as (pitch change-control), suggesting overlap between pitch height and visuospatial height processing. Crucially, this overlap is not likely due to some general sensitivity to changing stimuli since tactile processing differed from this pattern. Whereas tactile change resulted in a significant decrease in the BA17 ROI, there was no effect for tactile (height change-control), indicating that activity in this ROI was selective for changes in visuo-spatial height and auditory pitch.

General Discussion

Does processing pitch "height" activate areas of the cerebral cortex that are involved in processing spatial height? And if so, are they modality-specific or multimodal areas? These data provide preliminary answers to these questions. By comparing stimuli that differ in spatial height to those that remain at a constant position (control), we found activations in primary visual cortex (BA 17), an area shown previously to be sensitive to changes in spatial position (e.g., Bosking et al., 2002). We used this area as a visuospatial ROI in which to search for pitch-related activity. Crucially, we observed activity correlated with pitch in this primary visual region, suggesting that musical pitch may rely, in part, on unimodal visuospatial representations.

This is the first demonstration of overlap between processing of visuospatial height and pitch height in an ROI analysis, but more general activation of primary and secondary visual areas during pitch processing has been observed previously (e.g., Degerman et al., 2006; Foster & Zatorre, 2010; Platel et al., 1997). Taken together, these findings support the presence of modality-specific visuospatial activity during pitch processing. Beyond showing a general link between vision and audition (e.g., Romei, Murray, Cappe, & Thut, 2009), we find a cluster in BA 17 that is selective for processing changes in visuo-spatial height (as opposed to other aspects of visual stimuli) and is also selective for processing changes in pitch height (as opposed to other aspects of auditory stimuli).

Are multimodal spatial areas also involved in pitch processing? Our data provide no evidence for such involvement, but they do not rule out this possibility. Other studies have shown the IPL to be involved in pitch memory (Rinne, Koistinen, Salonen & Alho, 2009) and pitch production (Peck et al., 2009), and the IPS to be involved in pitch identification (Schwenzer & Mathiak, 2011) and pitch transformation tasks (Zatorre et al., 2010). Presumably, mechanisms that underlie pitch processing may differ depending on the complexity of the task. Our simple pitch comparisons differ from melody transformations (Zatorre et al., 2010) or complex pitch memory tasks (Rinne et al., 2009). These studies showing parietal activity for pitch processing did not directly compare space and pitch; it is possible that analyses like ours would reveal overlap between pitch and space in multimodal cortical areas in more complex pitch (and space) judgment tasks.

The present data provide initial evidence that a modality-specific brain area supports a link between the metaphorical "source domain" of space and "target domain" of pitch. This finding is notable given that, in general, evidence for modality-specific activity corresponding to metaphorical source domains like space has been elusive (see Willems & Casasanto, 2011, for review). An fMRI study by Quadflieg et al. (2011), for instance, confirmed that spatial height representations underlie representations of emotional valence. Yet, since patterns of co-activation were restricted to multimodal areas (i.e., IPL), Quadflieg et al.'s results do not provide any evidence that modality-specific activity underlies metaphorical mental representations.

By contrast, our results suggest that judgments of musical pitch depend in part on visual areas that are involved in spatial height processing. Although further studies are

needed to test for causal relationships between visual cortex activity and representations of space and pitch, modality-specific representations of spatial height may contribute to musical pitch processing, confirming a core assumption of embodied theories of metaphor.

References

Amorapanth, P. X., Widick, P., & Chatterjee, A. (2010). The neural basis for spatial relations. *Journal of Cognitive Neuroscience*, 22(8), 1739–1753.

Barsalou, L. W., Simmons, K. W., Barbey, A. K., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences*, 7(2), 84–91.

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

Bosking, W. H., Crowley, J. C., & Fitzpatrick, D. (2002). Spatial coding of position and orientation in primary visual cortex. *Nature Neuroscience*, 5(9), 874–882.

Casasanto, D. (2008). Similarity and Proximity: When does close in space mean close in mind? *Memory & Cognition, 36*(6), 1047-1056.

Degerman, A., Rinne, T., Salmi, J., Salonen, O., & Alho, K. (2006). Selective attention to sound location or pitch studied with fMRI. *Brain Research*, 1077(1), 123–134.

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.), *Proceedings of the 34th Annual Meeting of the Cognitive Science Society* (pp. 306–311). Austin, TX: Cognitive Science Society.

Eitan, Z., Ornoy, E., & Granot, R. Y. (2012). Listening in the dark: Congenital and early blindness and cross-domain mappings in music. *Psychomusicology: Music, Mind, and Brain*, 22(1), 33–45.

Evans, K. K., & Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features. *Journal of Vision*, 10(1), 6–12.

Foster, N. E. V., & Zatorre, R. J. (2010). Cortical structure predicts success in performing musical transformation judgments. *NeuroImage*, 53(1), 26–36.

Foster, N. E. V., Halpern, A. R., & Zatorre, R. J. (2013). Common parietal activation in musical mental transformations across pitch and time. *NeuroImage*, 75, 27–35.

Friston, K. J., Holmes, A., Poline, J. B., Price, C. J., & Frith, C. D. (1996). Detecting activations in PET and fMRI: levels of inference and power. *Neuroimage*, 4(3), 223–235.

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(3), 337–358.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by* (Vol. 111). Chicago London: The University of Chicago Press.

Lidji, P., Kolinsky, R., Lochy, A., & Morais, J. (2007). Spatial associations for musical stimuli: A piano in the head? *Journal of Experimental Psychology: Human Perception and Performance*, 33(5), 1189–1207.

Macaluso, E., & Driver, J. (2001). Spatial attention and crossmodal interactions between vision and touch. *Neuropsychologia*, 39(12), 1304–1316.

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

Meier, B. P., & Robinson, M. D. (2004). Why the Sunny Side Is Up Associations Between Affect and Vertical Position. *Psychological Science*, 15(4), 243–247.

Melara, R. D., & O'Brien, T. P. (1987). Interaction between synesthetically corresponding dimensions. *Journal of Experimental Psychology: General*, 116(4), 323–336.

Noesselt, T., Hillyard, S. A., Woldorff, M. G., Schoenfeld, A., Hagner, T., Jäncke, L., … Heinze, H. J. (2002). Delayed striate cortical activation during spatial attention. *Neuron*, 35(3), 575–587.

Peck, K. K., Galgano, J. F., Branski, R. C., Bogomolny, D., Ho, M., Holodny, A. I., & Kraus, D. H. (2009). Event-related functional MRI investigation of vocal pitch variation. *Neuroimage*, 44(1), 175–181.

Perry, D. W., Zatorre, R. J., Petrides, M., Alivisatos, B., Meyer, E., & Evans, A. C. (1999). Localization of cerebral activity during simple singing. *Neuroreport*, 10(18), 3979–3984.

Platel, H., Price, C., Baron, J. C., Wise, R., Lambert, J., Frackowiak, R. S., … Eustache, F. (1997). The structural components of music perception. A functional anatomical study. *Brain*, 120(2), 229–243.

Poline, J. B., Worsley, K. J., Evans, A. C., & Friston, K. J. (1997). Combining spatial extent and peak intensity to test for activations in functional imaging. *Neuroimage*, 5(2), 83–96.

Quadflieg, S., Etzel, J. A., Gazzola, V., Keysers, C., Schubert, T. W., Waiter, G. D., & Macrae, C. N. (2011). Puddles, parties, and professors: Linking word categorization to neural patterns of visuospatial coding. *Journal of Cognitive Neuroscience*, 23(10), 2636–2649.

Rinne, T., Koistinen, S., Salonen, O., & Alho, K. (2009). Task-dependent activations of human auditory cortex during pitch discrimination and pitch memory tasks. *The Journal of Neuroscience*, 29(42), 13338–13343.

Romei, V., Murray, M. M., Cappe, C., & Thut, G. (2009). Preperceptual and stimulus-selective enhancement of low-level human visual cortex excitability by sounds. *Current Biology*, 19(21), 1799–1805.

Rusconi, E., Kwan, B., Giordano, B. L., Umilta, C., & Butterworth, B. (2006). Spatial representation of pitch height: the SMARC effect. *Cognition*, 99(2), 113–129.

Saygin, A. P., McCullough, S., Alac, M., & Emmorey, K. (2010). Modulation of BOLD response in motion-sensitive lateral temporal cortex by real and fictive motion sentences. *Journal of Cognitive Neuroscience*, 22(11), 2480–2490.

Schubert, T. W. (2005). Your highness: vertical positions as perceptual symbols of power. *Journal of Personality and Social Psychology*, 89(1), 1–21.

Schwenzer, M., & Mathiak, K. (2011). Numeric aspects in pitch identification: an fMRI study. *BMC Neuroscience*, 12(1), 26–35.

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(1), 21–25.

Willems, R. M., & Casasanto, D. (2011). Flexibility in embodied language understanding. *Frontiers in Psychology*, 2.

Zatorre, R. J., Halpern, A. R., & Bouffard, M. (2010). Mental reversal of imagined melodies: a role for the posterior parietal cortex. *Journal of Cognitive Neuroscience*, 22(4), 775–789.

Zatorre, R. J., Meyer, E., Gjedde, A., & Evans, A. C. (1996). PET studies of phonetic processing of speech: review, replication, and reanalysis. *Cerebral Cortex*, 6(1), 21–30.