

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/116516>

Please be advised that this information was generated on 2016-05-02 and may be subject to change.

High pitches and thick voices:

The role of language in space-pitch associations

© Sarah Dolscheid, 2013

The research reported in this thesis was carried out at the Max Planck Institute for Psycholinguistics, the Donders Institute for Brain, Cognition and Behaviour, Center for Cognitive Neuroimaging, Radboud University Nijmegen, and the Baby Research Center Nijmegen. The research was supported by an IMPRS fellowship funded by the Max Planck Gesellschaft, München, Germany.

ISBN: 978-90-76203-51-5

Printed and bound by Ipskamp Drukkers b.v.

High pitches and thick voices:
The role of language in space-pitch associations

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. mr. S.C.J. Kortmann,
volgens besluit van het college van decanen
in het openbaar te verdedigen op dinsdag 22 oktober 2013
om 12.30 uur precies

door

Sarah Julia Dolscheid
geboren op 15 juli 1983
te Greven, Duitsland

Promotoren:

Prof. dr. Peter Hagoort

Prof. dr. Asifa Majid

Copromotor:

Dr. Daniel Casasanto (The University of Chicago, US)

Manuscriptcommissie:

Prof. dr. Ton Dijkstra

Prof. dr. Rolf Zwaan (Erasmus University Rotterdam)

Dr. Andrea Bender (University of Freiburg, DE)

Table of contents

Chapter 1	Introduction	1
Chapter 2	The thickness of musical pitch: Psychophysical evidence for linguistic relativity	15
Chapter 3	The sound of thickness: Prelinguistic infants' associations of space and pitch	35
Chapter 4	Space and pitch in language and thought: Exploring the locus of metaphor	51
Chapter 5	Searching high and low: The relation of space and musical pitch in the brain	67
Chapter 6	General Discussion and Conclusions	91
	References	105
	Samenvatting	119
	Acknowledgments	125
	Curriculum Vitae	129
	List of Publications	129
	MPI Series in Psycholinguistics	131

„Wird zugegeben, dass die Ausdrücke hoch und tief bei Tönen stets entsprechende Raumvorstellungen mit sich führen, so kann weiter gefragt werden, ob auch Töne, deren Auffassung augenblicklich nicht sprachlich eingekleidet wird, uns nichtsdestoweniger im räumlichen Kleide erscheinen; ob also immer, wenn wir Töne hören, sprach- und gedankenlos ihnen hingeben, Bilder des räumlich Tiefen, Hohen, Auf- und Absteigenden von selbst sich damit verknüpfen.“

(Stumpf, 1883, pp. 200-201)

Introduction

Chapter 1

Are *high* pitches somewhere in the air? And can you feel *warm* colors? Or taste a *sweet* voice? Although these questions may seem exceptional, expressions like a 'high pitch' are very common in language. People often talk about one domain metaphorically, in terms of another. For instance, English speakers frequently talk about time using spatial language (e.g., a *long* journey, a *short* meeting) or about valence by using expressions of upward versus downward motion or position in space (e.g., her spirits *soared* vs. he is feeling *low*). In these classic cases, people recruit metaphors from more concrete or perceptually rich domains to talk about rather 'abstract' target domains (Clark, 1973; Lakoff & Johnson 1980, 1999). The space-time relation is a typical example: While space can be perceived by vision, touch, or movement, the domain of time is argued to be more abstract as people can only imagine the temporal (see for instance Clark, 1973).

There are, however, also conventional linguistic metaphors for domains people can perceive directly. Sensory impressions themselves are often described metaphorically, in terms of other modalities or spatial dimensions. As illustrated above, people use so-called cross-modal (Marks, 2011) or synesthetic (Yu, 2003) metaphors in which colors can be *loud*, smells can be *sharp*, and pitches can be *high*. This linguistic 'mingling' of different sensory modalities has led a number of researchers to suggest that cross-modal metaphors may reveal something about underlying modes of cross-modal perception (e.g., Eagleman & Cytowic, 2009; Marks, 1982, 1987; Williams, 1976). A related question has been debated extensively in the metaphor literature: Do people only talk in terms of metaphors, or do they also *think* metaphorically (e.g., Lakoff & Johnson, 1980; McGlone, 2007; Murphy 1996, 1997)? In their seminal book, "Metaphors we live by", Lakoff and Johnson (1980) posited that metaphors are not merely linguistic devices but that they reveal important aspects of cognition. Whereas support in favor of this proposal was initially linguistic (Lakoff & Johnson, 1980) and thus

circular in nature (McGlone, 2007), there is now rich experimental evidence suggesting metaphorical thinking in domains like time (Boroditsky, 2000), number (Dehaene, Bossini, & Giraux, 1993), or preference (Casasanto, 2008). People thus seem to think metaphorically, just like patterns in language suggest.

However, little is known about the actual role that metaphoric language plays with respect to nonlinguistic cognition. In my thesis, I will examine this very role of language by focusing on cross-modal metaphors. Since these metaphors predominantly occur in the domain of sound (Day, 1996; Ullmann, 1963; Williams, 1976), the focus will be on the auditory modality. More specifically, I will zoom in on spatial metaphors of auditory pitch.

In English and many other languages, the pitch of a tone can be described spatially, as being *high* or *low* (e.g., Stumpf, 2006). Moreover, there is evidence that pitch and space are importantly related in people's minds (e.g., Melara & O'Brien, 1987; Evans & Treisman, 2010). This observation raises a question that is essential to my dissertation: What is the relationship between space-pitch metaphors in language and in thought? More specifically, I am interested in the role that language plays with respect to space-pitch associations. Does language merely 'reflect' nonlinguistic space-pitch associations or can language also influence space-pitch mappings? Do metaphors in language establish associations between space and musical pitch in people's minds? What are the neuronal underpinnings of space-pitch metaphors?

In order to address these questions, I combine a range of methods including cross-linguistic observations, psychophysical experiments, infant studies and neuroimaging. Before introducing my research questions in more detail, I will first define musical pitch.

Musical pitch

Musical or auditory pitch is a psychoacoustic property that approximately corresponds to the frequency of a sound (e.g., Moore, 1994). Whereas pitch and frequency are closely related, they are not equivalent. Pitch is a subjective attribute of a sound stimulus, and as such cannot be measured directly. Still,

pitch is usually quantified as frequency in cycles per second (hertz) (e.g., Moore, 1994).

While a pure tone consists of a single sinusoidal waveform whose frequency can be easily determined, pitches of complex tones consist of multiple sinusoids superimposed (e.g., Walker, Bizley, King, & Schnupp, 2010). Sinusoids that are (integer) multiples of one another are called harmonics. The pitch of a complex tone is roughly equivalent to the highest common divisor of the sound's harmonics, which in turn is called the fundamental frequency (f_0). In most naturally occurring sounds the fundamental frequency (f_0) corresponds to the lowest harmonic that is present in the sound. However, sometimes the perception of pitch can persist at the lowest harmonic even if this very frequency is absent from the sound (missing fundamental) (Walker et al., 2010). Pitch (f_0) perception can thus be maintained as long as the relation between the remaining harmonics is unchanged.

People perceive pitches within the range of approximately 30 to 5000 hertz (Bizley & Walker, 2010) with sounds of higher frequencies being perceived as 'higher' in pitch (Moore, 1994; Walker et al., 2010). Usually, people are able to discriminate very fine changes in pitch (Sano & Jenkins, 1991). However, the just noticeable difference (JND; the threshold at which a change in pitch is perceived) varies across people and over the range of hearing. For intermediate frequencies (between 500 and 2000 hertz) the JND is around 0.3%, indicating that people can discriminate two pitches when they differ by 0.3% (Sano & Jenkins, 1991).

Pitch is not a unidimensional attribute that corresponds solely to frequency, but rather it consists of two dimensions (Hubbard, 1996; Shepard, 1982; Ueda & Ohgushi, 1987): Pitch 'height' (absolute frequency) and tone chroma (an attribute describing similarity of pitch quality). For example, the 'middle C' (C4) and higher Cs (C5, C6, etc.) on a Western musical scale differ in pitch height. Yet they all collectively describe a single pitch class and therefore share the same pitch chroma.

Whereas pitch chroma indicates that pitch perception is to some extent periodic (C4 and C5 sound similar, although they are an octave apart), pitch height is defined as "that auditory attribute of sound according to which sounds can be ordered on a scale from low-to-high" (American National Standards

Institute, 1994, cited in Walker et al., 2010). To put it simply, variations in pitch height "give rise to a sense of melody" (Moore, 2010, p.472).

Language of Pitch

As is clear from the preceding passage, it is almost impossible to define pitch without using spatial terminology such as 'pitch height' (at least in English). In fact, many languages encode pitch in terms of vertical space. As early as 1883, the German philosopher and psychologist Carl Stumpf collected pitch vocabulary from a number of languages including Sanskrit, Latin, and Hebrew. Stumpf found that most of the languages in his sample applied the labels high and low to the pitches of tones (Stumpf, 2006). Likewise languages like English, Dutch, German, Spanish, Italian, Greek, Polish, Danish and Icelandic, but also non-Indo-European languages like Japanese, Chinese and Siwu, a Kwa language spoken in Ghana, encode pitch in terms of spatial height (see e.g., Dingemanse, 2011; Rusconi, Kwan, Giordano, Umiltà, & Butterworth, 2006). This finding has led some researchers to the conclusion that height-pitch metaphors may be universal (e.g., Evans & Treisman, 2010; Pratt, 1930).

However, not all languages metaphorize pitch spatially. In Kreung, an Austro-Asiatic language spoken in Cambodia (Lewis, 2009), high and low pitches are described as 'tight' and 'loose' (Parkinson, Kohler, Sievers, & Wheatley, 2012). The Kpelle people of Liberia talk about high and low pitches as 'light' and 'heavy', and the Bashi people of central Africa call high and low pitches 'weak' and 'strong' (Eitan & Timmers, 2010). Even languages that encode pitch spatially do not necessarily use the same height-pitch metaphors that are familiar to English speakers. There is evidence suggesting that pitch metaphors of ancient Greek speakers used to follow a reversed pattern with high referring to low pitches and vice versa (Stumpf, 2006). Furthermore, for the Manza of Central Africa, high pitches are 'small' and low pitches 'large' (Stone, 1981) and in languages like Farsi, Turkish, and Zapotec (spoken in the Sierra Sur of Mexico) high pitches are referred to as 'thin' and low pitches as 'thick' (Shayan, Ozturk, & Sicoli, 2011).

Pitches are thus not necessarily described in terms of vertical space.¹ But whereas Miller and Johnson-Laird (1976) considered it "ironic that people use vocal sounds to name everything else yet have such a limited vocabulary of names for sounds themselves" (p. 25), it seems a common cross-linguistic strategy to overcome this terminological gap by using metaphoric language, especially in the domain of musical pitch (see Dolscheid, 2008; Ullmann, 1963; Williams, 1976).

Metaphors in language = metaphors in thought?

Can metaphors in language reveal something about thought? This question has been debated to some extent in the metaphor literature (Lakoff & Johnson, 1980; McGlone, 2007; Murphy 1996, 1997). There is now evidence that people think metaphorically in domains like power (Schubert, 2005), number (Dehaene, Bossini, & Giraux, 1993) or time (Casasanto & Boroditsky, 2008). Speakers of English, for instance, not only conventionally use distance metaphors to talk about duration (e.g., a 'long' time) but they also represent time in terms of spatial extent (distance) in completely nonlinguistic tasks (Casasanto & Boroditsky, 2008). This finding suggests that spatial metaphors of time are more than mere linguistic conventions.

What then is the case for metaphors of musical pitch? The height-pitch metaphor that is familiar to speakers of many languages seems to be highly conventionalized. In fact, it seems to be conventionalized to the extent that some researchers have called it a "dead" metaphor (Dempsey, 2008, p.8). But are height-pitch metaphors really 'dead'? Or can they reveal something fundamental about how people think about musical pitch? Ideas about close links between physical space and musical space have been described in some of the earliest works of Greek musical theory (e.g., by Aristoxenos) (Barker, 1987; Cupchik,

¹ Note that Carl Stumpf, although often cited as having claimed that all of the world's languages apply the labels "high" and "low" to describe pitches (e.g., Evans & Treisman, 2010; Pratt 1930; Spence, 2011), acknowledged that not all languages use height-pitch metaphors (Stumpf, 2006, p. 192).

Phillips, & Hill, 2001). Wilhelm Wundt (1893) as well as Carl Stumpf (2006) speculated about height-pitch associations and their role beyond linguistic expressions. The first to empirically test height-pitch associations, however, was Carroll Pratt (1930). In a seminal study, participants were asked to locate the position of different pitches on a numbered scale. Tones were emitted by a hidden loudspeaker that varied in location. Independent from actual spatial variation, participants consistently ascribed lower pitches to lower positions in space and conversely higher pitches to higher positions, suggesting that people also represent pitch in terms of spatial height (Pratt, 1930). This finding has been confirmed in more recent speeded classification tasks. People have been found to respond faster to a stimulus when visuospatial height and pitch correspond compared to incongruent stimulus presentation (Melara & O'Brien, 1987; see also Ben-Artzi & Marks, 1995; Evans & Treisman, 2010). Participants are also faster to press response keys located higher in space for high-frequency pitches than for low-frequency pitches, and vice versa, as was demonstrated in stimulus-response compatibility experiments (Lidji, Kolinsky, Lochy, & Morais, 2007; Rusconi et al., 2006). Beyond experimental evidence, height-space relations also find their expression in Western musical notation, with higher pitches written higher on the musical staff.

It thus seems that speakers of many languages not only talk about pitch in terms of spatial height but they also represent pitches as being high or low in space. What then is the relationship between space-pitch metaphors in language and (nonlinguistic) space-pitch associations in thought?

The role of language in space-pitch associations

There are varying positions concerning the relationship of space-pitch mappings and linguistic metaphors. On the one hand, it has been argued that space-pitch associations may exist independently of language. In George Lakoff's words: "[T]he locus of metaphor is not in language, at all, but in the way we conceptualize one mental domain in terms of another" (Lakoff, 1993, p. 203). Support for this idea comes from experiments with prelinguistic infants. It has

been shown that associations between verticality and pitch are present early in life, prior to the acquisition of language. In a preferential looking task, 3- to 4-month-old infants look longer at a ball moving upwards if it is accompanied by a rising pitch than if it is accompanied by a falling pitch (Walker et al., 2010). Also, prelinguistic infants were found to 'match' visual arrows pointing up or down with tones sweeping up or down in frequency (Wagner, Winner, Cicchetti, & Gardner, 1981). These findings suggest that language may simply 'echo' (pre-) existing space-pitch associations. Or as Lakoff (1993) pointed out: The metaphorical expressions in language may merely be "surface realization[s]" of underlying "cross-domain mappings" (p. 203). If this were true, linguistic metaphors should not have an impact on nonlinguistic space-pitch associations.

Contrary to this position, the co-occurrence of linguistic height-pitch metaphors and height-pitch relations has led other researchers to suggest that associations between spatial height and pitch may be influenced by language (e.g., Martino & Marks, 1999; Melara & Marks, 1990; see also Spence, 2011). Over the course of language use, height-pitch associations are proposed to become mediated by "linguistic or semantic systems" (Martino & Marks, 1999, p. 904). Metaphors in language would thus not merely echo metaphors in thought but rather influence these associations. In this case linguistic patterns could affect nonlinguistic cognition.

One main goal of this thesis is to find out which role language actually plays. To what extent are metaphors in language and space-pitch associations linked? Does language merely reflect space-pitch associations? Or do metaphors in language have an impact on space-pitch mappings? Cross-linguistic diversity can help us to tackle these questions. Since not everybody uses the same metaphors for pitch in language one can test whether people who use different linguistic metaphors represent pitch differently, according to the languages they speak.

Question 1: Can language have an influence on nonlinguistic pitch representations?

This question of linguistic relativity, often associated with the writings of Benjamin Whorf (1956), has been extensively debated in domains like 'time' (e.g., Boroditsky, 2001; Casasanto, 2008; Chen, 2007), 'space' (e.g., Majid, Bowerman, Kita, Haun, & Levinson, 2004; Li, Abarbanell, Gleitman, & Papafragou, 2011), 'motion' (e.g., Gennari, Sloman, Malt, & Fitch, 2002; Papafragou, Hulbert, & Trueswell, 2008), and 'color' (e.g., Heider & Olivier, 1972; Regier & Kay, 2009), but hardly anything is known about effects of language on pitch representation.

In Chapter 2, I therefore investigate the influence of different linguistic pitch metaphors on nonlinguistic cognition, by testing speakers of Dutch and Farsi. Whereas Dutch speakers describe the pitch of a sound as 'high' (hoog) or 'low' (laag), Farsi speakers describe high-frequency tones as 'thin' (nāzok) and low-frequency tones as 'thick' (koloft). In order to find out about Dutch and Farsi speakers' pitch representations, my colleagues and I conducted a pair of non-linguistic psychophysical space-pitch interference experiments. Participants were asked to reproduce musical pitches that they heard in the presence of irrelevant spatial information (i.e. lines that varied either in height or in thickness). If Dutch and Farsi speakers mentally represent pitch the way they talk about it, using different kinds of spatial representations, they should show contrasting patterns of cross-dimensional interference: Dutch speakers' pitch estimates should be more strongly affected by irrelevant height information, and Farsi speakers' by irrelevant thickness information. If language, however, does not have an impact on nonlinguistic pitch representations, Dutch and Farsi speakers should show similar patterns of spatial interference. Results reveal that Dutch and Farsi speakers differ in their patterns of space-pitch interference, suggesting that metaphors in language can indeed influence nonlinguistic pitch representations.

Question 2: How might language influence nonlinguistic pitch representations?

While linguistic metaphors appear to influence nonlinguistic space-pitch associations, different mechanisms of how language affects thought could have been at work. According to one influential view of the relationship between language and thought, patterns in language can influence nonlinguistic mental representations while people are packaging their thoughts into words (Slobin, 1996). Similarly, language may influence thought while speakers perform tasks for which verbal codes can be helpful (e.g., Gennari, et al., 2002; Papafragou et al., 2008). Supporting this view, previous studies have found that some linguistic relativity effects disappear under the influence of competing verbal interference (i.e. overt or covert use of irrelevant language during the experimental task; Roberson & Davidoff, 2000; Wiggett & Davies, 2008; Winawer et al., 2007). In the locus classicus of linguistic relativity – the color domain – it has been found that cross-linguistic differences in color vocabularies can result in differences in color categorization, color memory and even color perception (Davidoff et al., 1999; Roberson, Davidoff, Davies, & Shapiro, 2005; Roberson, Pak, & Hanley, 2008; Winawer et al., 2007; Thierry, Athanasopoulos, Wiggett, Dering, & Kuipers, 2009). English and Russian color terms, for instance, divide the color spectrum differently. Unlike English, Russian makes an obligatory distinction between darker blues ('siniy') and lighter blues ('goluboy'). Russian speakers are also faster to discriminate two colors when those colors fall into different linguistic categories in Russian (one siniy and the other goluboy) than when they are from the same linguistic category (both siniy or both goluboy) in a speeded color discrimination task (Winawer et al., 2007). English speakers, however, do not show any category advantage. Interestingly, the advantage of between category discrimination (siniy vs. goluboy) in Russian speakers disappears under verbal interference (silent rehearsing of number series during the color task). In these and related experiments, it has been argued that the 'online' use of language drives linguistic relativity effects and can therefore be disrupted by distracting verbal tasks (e.g., Winawer et al., 2007; for a critical perspective see Lupyan, 2012).

In chapter 2 I test whether space-pitch associations are also mediated by the online use of language (as would also be suggested if height-pitch associations were directly mediated by semantic systems; e.g., Martino & Marks, 1999). Therefore, in addition to the use of nonlinguistic psychophysical tasks, I test Dutch speakers' height-pitch associations while keeping them busy with the rehearsal of letter strings. If height-pitch associations are mediated by the online use of language, effects of spatial height on pitch should disappear under verbal interference. On the contrary, if height-pitch associations are not mediated by the online use of language, effects of spatial height on pitch should persist under verbal interference. Unlike other linguistic relativity effect, my results show that height-pitch associations persisted under verbal interference, suggesting that language does not affect pitch representations in an online fashion.

Question 3: Does language establish associations between space and pitch in the first place?

In chapter 2 I show that language can influence nonlinguistic space-pitch representations. However, it remains unclear whether language establishes links between space and pitch in the first place or whether it merely modulates space-pitch associations. As previous findings suggest, height-pitch associations appear to be established prior to language (Walker et al., 2010; Wagner et al., 1981). Critically, however, nothing is known about the origins of other space-pitch relationships, such as the association between thickness and pitch. Do children also have this mapping available to them prior to language? Or is it only learned on the basis of language-input? On one possibility, children start out with both a height-pitch and a thickness-pitch mapping even before they learn language. On the other hand, height-pitch and thickness-pitch associations might follow different trajectories. Whereas height-pitch associations appear to be available to prelinguistic infants, the thickness-pitch mapping might only be learned later. (e.g., on the basis of metaphors in language; Gentner, 2003; Martino & Marks, 1999). In order to determine the prelinguistic availability of space-pitch mappings, I tested 4-month-old Dutch babies in a preferential looking paradigm

(chapter 3). To investigate height-pitch correspondences, infants watched a ball moving up and down the screen accompanied by the sound of a sliding whistle (based on Walker et al., 2010). The sound was either congruent or incongruent with the ball's movement. In an analogous thickness-pitch task infants were presented with the same sound but simultaneously watched a vertical tube changing in thickness. I reasoned that if both height-pitch and thickness-pitch mappings are available to infants prelinguistically, infants should prefer congruent height-pitch as well as congruent thickness-pitch stimuli over incongruent ones. On the other hand, if height-pitch and thickness-pitch relationships follow different developmental trajectories, with thickness mappings only becoming acquired later, then infants should prefer congruent height-pitch stimuli but not congruent thickness-pitch stimuli. My results reveal that infants start out with both a thickness-pitch and a height-pitch mapping available prior to language acquisition. These findings suggest that language is unlikely to establish associations between space and pitch in the first place.

Question 4: The locus of space-pitch metaphors – is it in language or in thought?

While language can shape space-pitch associations, it does not seem to instill these mappings on prelinguistic infants (chapter 3). Moreover, space-pitch associations of adult speakers remain unaffected by online language use (chapter 2). It thus seems that language may not be the main 'ingredient' in space-pitch metaphors. Rather, in line with conceptual metaphor theory, linguistic metaphors like 'a high pitch' may reflect the activation of a particular non-linguistic spatial schema in a person's mind. The spatial schema itself may thus be more important than its linguistic form. Following this argument, it should be possible to activate mental metaphors using the wrong words, so long as the words activate the right schema. To test this prediction, English speakers were tested in a (Stroop-like) space-pitch congruity task. Participants either classified high- and low-frequency pitches using the conventional terms 'high' and 'low' or using the words 'tall' and 'short', which also refer to the upper and lower poles of a vertical spatial continuum. Participants were faster to classify pitches in a

congruent way (i.e. high pitches as 'high' and low pitches as 'low') than to classify pitches in an incongruent fashion (i.e. high pitches as 'low' and low pitches as 'high'). Moreover, a similar congruity effect was found when participants classified pitches as 'tall' and 'short', even though these words are not conventionally used in English to refer to pitch (e.g., a tall pitch does not mean a high pitch). A second experiment revealed that this effect was specific to spatial words that cued a spatial schema of verticality, and that it was not an artifact of linguistic markedness or polarity alignment. Results indicate that mental metaphors for pitch can be activated by lexically inappropriate words so long as they cue the right spatial schema. These findings suggest that the locus of height-pitch metaphors is in thought rather than in language.

Question 5: How are space-pitch associations represented in the brain?

To further explore the nature of space-pitch mappings and their instantiations in the brain, I investigated possible spatial bases that may underlie representations of musical pitch (chapter 5). For this purpose I used functional Magnetic Resonance Imaging (fMRI). In fMRI the level of blood oxygenation is measured in the brain. The blood-oxygenation-level dependent (BOLD) contrast provides an indicator of which regions in the brain areas are more active during a task. Although the temporal resolution of fMRI is quite low (~ 6 s), it can localize very precisely where something is happening in the brain (~ 3 mm resolution) (Huettel, Song, & McCarthy, 2004). This method thus provides a helpful tool in localizing shared neural circuitry related to space and pitch processing. In order to localize spatial activations, participants completed spatial judgment tasks in two modalities: vision and touch. Moreover, they compared tones that varied in pitch. I reasoned that if pitch judgments recruit the same neural circuitry as judgments regarding visual space, tactile space, or the combination of the two (i.e. multimodal) space, one should see pitch activations in the respective region(s) that are activated by spatial height processing. Using fMRI thus provided an opportunity to distinguish between unimodal and

multimodal spatial bases of musical pitch, contributing further information about the exact nature of space-pitch associations.

Summary

This thesis seeks to explore space-pitch metaphors in language and in thought. It particularly focuses on the role of language in space-pitch associations by combining cross-linguistic approaches, psychophysical methods, infant studies and neuroimaging. Chapter 2 lays the groundwork by determining whether and how language can have an impact on nonlinguistic pitch representations. In chapter 3, we look at space-pitch associations in prelinguistic infants. Finally, in chapters 4 and 5, the focus is on the nature of space-pitch associations. Where is the locus of space-pitch associations (chapter 4) and to what extent do we find modality-specific height-pitch representations in the brain (chapter 5)?

The thickness of musical pitch: Psychophysical evidence for linguistic relativity

Chapter 2

Based on:

Dolscheid, S., Shayan, S., Majid, A., & Casasanto, D. (2013). The thickness of musical pitch: Psychophysical evidence for linguistic relativity. *Psychological Science*, 24(5), 613-621.

Abstract

Do people who speak different languages think differently, even when they are not using language? To find out, we compared mental representations of musical pitch in native speakers of Dutch and Farsi, using nonlinguistic psychophysical tasks. Whereas Dutch speakers describe pitches as 'high' (*hoog*) or 'low' (*laag*), Farsi speakers describe high-frequency pitch as 'thin' (*nāzok*) and low-frequency pitches as 'thick' (*koloft*). Differences in language were reflected in differences in performance on two pitch reproduction tasks, even though the tasks used simple nonlinguistic stimuli and responses. To test whether experience using language changes pitch representations, we trained native Dutch speakers to use Farsi-like metaphors, describing pitch in terms of thickness. After training, Dutch speakers' performance on a nonlinguistic psychophysical task resembled native Farsi speakers'. People who use different space-pitch metaphors in language also think about pitch differently. Language plays a causal role in shaping nonlinguistic representations of musical pitch.

Introduction

In English, musical pitches can be 'high' or 'low', melodic lines can 'rise' or 'fall' and people can sing at the 'top' or the 'bottom' of their range. Are these spatial metaphors merely linguistic conventions, or do they reflect something fundamental about the way people mentally represent musical pitch?

Space and pitch are importantly related in the brain and mind. Amusic patients, who have difficulty discriminating pitch changes, also show spatial deficits (Douglas & Bilkey, 2007; but see Tillmann et al., 2010). In spatial compatibility tasks, healthy participants are faster to press high response keys for high-frequency pitches than for low-frequency pitches, and vice versa for low response keys (Lidji, Kolinsky, Lochy, & Morais, 2007; Rusconi, Kwan, Giordano, Umiltà, & Butterworth, 2006). Beyond binary high-low correspondences, psychophysical pitch reproduction tasks show that pitch maps onto vertical space in a continuous analog fashion (Casasanto, 2010), as predicted by theories of metaphorical mental representation (Lakoff & Johnson, 1980).

Pitch and vertical space interact even in prelinguistic infants. In a preferential looking task, 3- to 4-month-olds preferred congruent trials (in which visuospatial height and pitch height corresponded) over incongruent trials (Walker et al., 2010; see also Wagner, Winner, Cicchetti, & Gardner, 1981). It appears that people mentally represent pitch in terms of vertical space, just like they talk about it in English and many other languages.

However, not everybody talks about pitch in the same way. The Kpelle people of Liberia, for instance, talk about high and low pitches as 'light' and 'heavy'. The Suyá people of the Amazon basin call high pitches 'young' and low pitches 'old', and the Bashi people of central Africa call high pitches 'weak' and low pitches 'strong' (Eitan & Timmers, 2010).

Even languages that use spatial metaphors for pitch may not use the same vertical metaphors that are familiar to English speakers. For the Manza of Central Africa, high pitches are 'small' and low pitches 'large' (Stone, 1981). In other languages like Farsi, Turkish, and Zapotec (spoken in Mexico) high pitches are 'thin' and low pitches 'thick' (Shayan, Ozturk, & Sicoli, 2011; Shayan, Ozturk, Bowerman, & Majid, submitted).

Do people who use different metaphors in language mentally represent pitch differently? If so, how deep are the effects of language on musical pitch? Could language shape the nonlinguistic representations that people use for perceiving or producing musical pitches, even when they are not using language?

A first hint that people who use different pitch metaphors think about pitch differently comes from co-speech gestures. Consistent with the Manzas' linguistic coding of pitches as small and large, speakers have been observed lowering their hand in space while referring to smaller (i.e., higher) pitches, contrary to the English high-low mapping (Ashley, 2004). This suggests that people may conceptualize pitch consistent with their pitch vocabulary. However, gestures that match the co-occurring speech may only reveal conventions for communicating about musical pitches, not modes of conceptualizing them. Alternatively, they may reveal a 'shallow' influence of language on thought, indicating that people do indeed conceptualize pitch in language-specific ways, but only while they are packaging their thoughts into words (i.e., while they are "thinking for speaking", see Slobin, 1996).

A persistent challenge in testing relationships between language and nonlinguistic mental representations is devising truly nonlinguistic tasks. Here we used a pair of psychophysical tasks with nonlinguistic stimuli and responses to test pitch representations in speakers of a language with 'height' metaphors (Dutch) and another with 'thickness' metaphors (Farsi). In one task (Height Interference), participants saw lines at varying heights while listening to tones of different pitches. After each tone, participants reproduced the pitch by singing it back. In the other task (Thickness Interference) participants saw lines of varying thickness while hearing tones of different pitches, and sang back the pitches that they heard. In both tasks, the spatial information was irrelevant, and spatial variation was orthogonal to variation in pitch. As such, the spatial dimension of the stimuli served as a distractor: a piece of information that could potentially interfere with performance on the pitch reproduction task.

We reasoned that if Dutch and Farsi speakers' mental representations of pitch were similar irrespective of the languages they speak, then performance on these tasks should not differ between language groups. On the other hand, if Dutch and Farsi speakers conceptualize pitch the way they talk about it,

activating different kinds of spatial representations, they should show contrasting patterns of cross-dimensional interference: Dutch speakers' pitch estimates should be more strongly affected by irrelevant height information, and Farsi speakers' by irrelevant thickness information.

Experiment 1: Do People Think About Pitch Like They Talk About It?

Methods

Participants.

Native Dutch speakers ($n=20$) and native Farsi speakers ($n=20$) performed the height-interference task. Likewise, native Dutch ($n=20$) and Farsi speakers ($n=20$) performed the thickness-interference task. Within each language group, some participants were tested in both experiments, with the order of tasks counterbalanced and the testing sessions separated by at least a week. One additional Farsi-speaking participant was tested in the height-interference, but was excluded for not following the instructions. The data of three additional Farsi-speaking participants could not be analyzed due to sound recording problems. Dutch participants were recruited from the Max Planck Institute participant pool. Farsi speakers were recruited from Nijmegen, Delft and Leiden.² All participants received payment to participate.

Materials.

For the Height Interference experiment (adapted from Casasanto, 2010), horizontal lines intersected a vertical reference line at one of nine different heights (ranging from 80 to 720 pixels from bottom to top of the computer screen, in 80 pixel increments). For the Thickness Interference experiment, a vertical line appeared in the middle of the screen in one of nine thicknesses (ranging from 8 to 72 pixels in 8 pixel increments). Variation in thickness was thus proportional to variation in height. In each experiment, the nine different

² All Farsi-speaking participants spoke Farsi on a daily basis, according to a language background questionnaire. Most also spoke some Dutch or English, and may have been exposed to height metaphors. In principle, this exposure could eliminate the predicted effects, but it could not produce them.

lines were fully crossed with nine different pitches ranging from C4 to G#4 in semitone increments, to produce 81 distinct trials. Stimuli were presented on a pc laptop (resolution = 1024x768 pixels) using Presentation software (www.neurobs.com). Lines were presented in white on a grey background (350 pixels wide). Auditory stimuli consisted of pure tones that were created using Audacity software (<http://audacity.sourceforge.net/>) and presented through sealed headphones at a constant amplitude for two seconds.

Singing responses were recorded by an EDIROL R-09 recording device, and analyzed using Praat software (<http://www.fon.hum.uva.nl/praat/>) by a coder blind to the corresponding spatial stimuli. The approximate temporal midpoint of each response was determined by visual inspection of the waveform. The average fundamental frequency (F_0) of each sung response was extracted from a 600 ms interval spanning 300 ms before and after the estimated temporal midpoint, to ensure that measured F_0 was representative of the whole response.

Written instructions were presented in the native language of the participant. They contained no space-pitch metaphors. The tasks, themselves, comprised only nonlinguistic stimuli (lines and tones) and nonlinguistic responses (sung tones).

Procedure.

Participants were asked to watch the lines and listen to the tones carefully, and to sing them back as accurately as possible.

After three practice trials, participants were presented the 81 line-pitch pairings one at a time, in random order, for 2 seconds each. Immediately after each stimulus, a picture of a microphone appeared in the center of the screen indicating that the participants had 2 seconds to sing back the pitch they had heard. Each response period was followed by an inter-trial interval of 500 milliseconds. After 40 trials, participants had a self-paced break. Testing lasted about 15 minutes, and was followed by a debriefing.

Results

Pitch estimation, cross-domain effects.

For each participant the values of the height or thickness stimuli were normalized, and we computed the slope of the effect of normalized spatial height or thickness on participants' reproduced pitches. In Dutch speakers, the spatial height of the stimuli influenced pitch estimates, according to a one-sample t-test comparing the mean of the normalized slopes against zero, $\text{Slope}=2.65$, $t(19)=2.70$, $p=.01$ (all tests two-tailed). Tones accompanied by higher lines were reproduced at a higher frequency on average than the same tones accompanied by lower lines. By contrast, the thickness of the stimuli had no significant effect on pitch reproduction, $\text{Slope}=0.60$, $t(19)=0.57$, *ns*. Farsi speakers showed the opposite pattern: Thickness influenced pitch estimates, as predicted by thickness metaphors in Farsi (i.e., tones accompanied by thicker lines were reproduced at a lower frequency on average than the same tones accompanied by thinner lines; $\text{Slope}=-2.85$, $t(19)=2.09$, $p=.05$, but height had no significant effect on pitch reproduction, $\text{Slope}=-0.71$, $t(19)=1.16$, *ns*.)

To test for the predicted interaction of Language (Dutch, Farsi) and Task (Height Interference, Thickness Interference), the normalized slopes from the thickness task were multiplied by -1. This multiplication was necessary because the relationship between spatial magnitude and frequency reverses between height and thickness metaphors: Greater spatial height corresponds to higher frequency, but greater spatial thickness corresponds to *lower* frequency. Multiplying the slopes by -1 for one of the tasks aligns the space and pitch continuums (i.e., the slope then indicates the same relationship between spatial magnitude and frequency for both tasks). According to a 2 x 2 ANOVA, Language interacted with Task to predict the effect of space on pitch estimates, $F(1,79)=10.73$, $MSE=.29$, $p=.002$, consistent with the use of height metaphors in Dutch and thickness metaphors in Farsi (Figure 1a, also see Figure 2). There were no main effects.

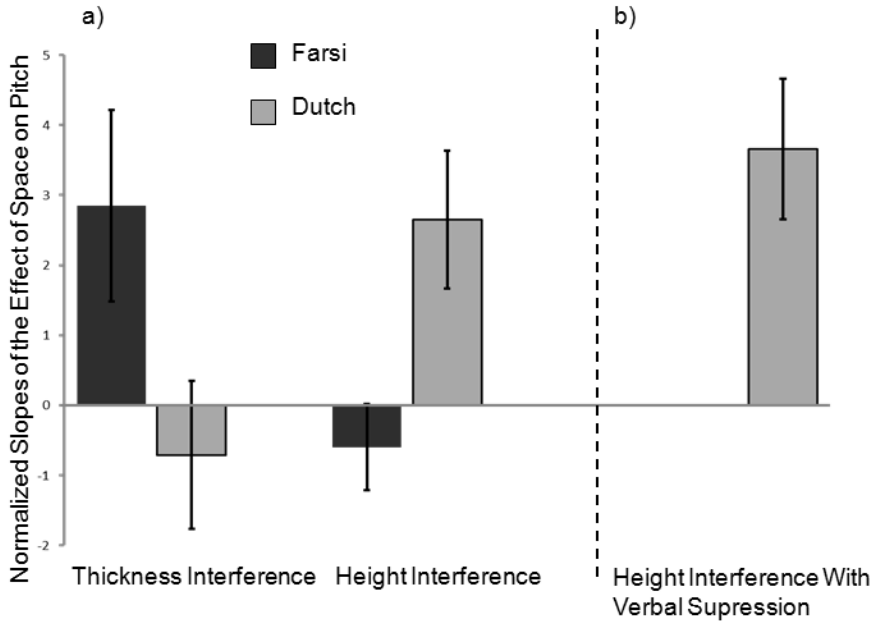


Figure 1. Results of Experiments 1 and 2. Normalized slopes of the effect of space on pitch reproduction are plotted as a function of spatial interference type. (a) Experiment 1: the influence of height and thickness information on pitch reproduction in Farsi and Dutch speakers. The effect of height interference was greater in Dutch speakers than in Farsi speakers, $t(38)=2.90$, $p=.01$ and conversely, the effect of thickness interference was greater in Farsi speakers than in Dutch speakers, $t(38)=2.00$, $p=.05$. In Dutch speakers the effect of height interference was greater than the effect of thickness interference, $t(38)=2.26$, $p=.03$, whereas for Farsi speakers the effect of thickness interference was greater than the effect of height interference, $t(38)=2.38$, $p=.02$. (b) Experiment 2: the influence of height information on pitch reproduction in Dutch speakers under verbal interference (Experiment 2). Error bars represent standard error of the mean.

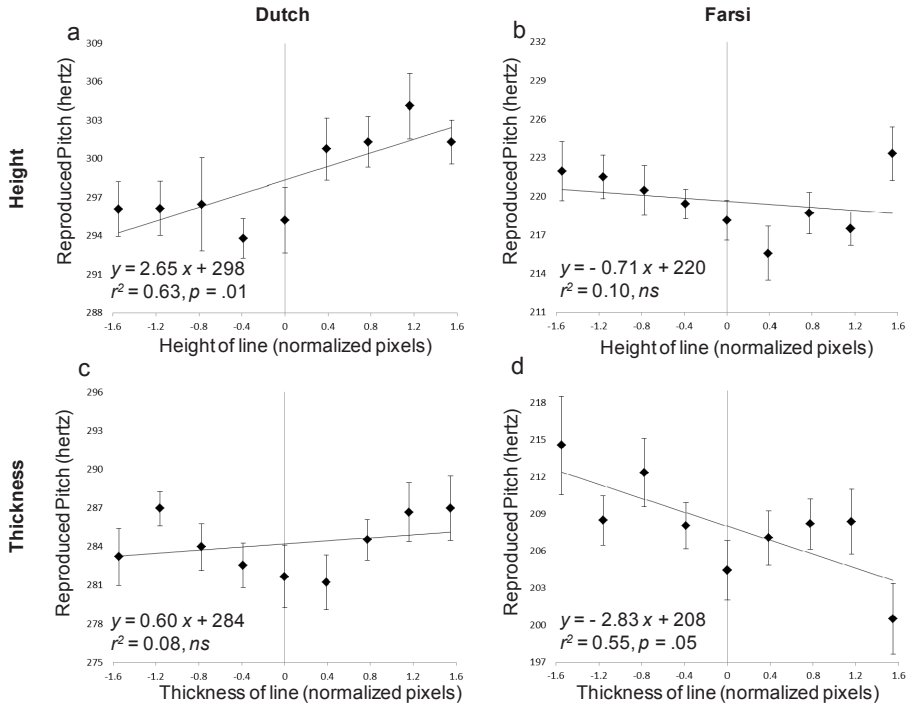


Figure 2. Results of Experiment 1. Effects of height interference (top) and thickness interference (bottom) on pitch estimates in speakers of Dutch (left) and Farsi (right). Error bars represent standard error of the mean.

Musical experience

In a further analysis, we tested whether differences in musical experience between Dutch and Farsi speakers affected performance. In addition to linguistic height-pitch metaphors in Dutch, there are also associations between height and pitch in musical notation, in both Dutch and Farsi culture: Higher on the staff is higher in pitch. Therefore it is possible that differences in experience reading musical notation could have contributed to the observed effect. During a debriefing, participants rated how well they read music on a scale of 1-7. Musical experience did not interact with Language and Task, $F(1,79)=1.52$, $MSE=.41$, ns , and the interaction between Language and Task remained highly significant

when the effect of musical experience was controlled, $F(1,79)=10.83$, $MSE=.29$, $p=.002$.

Pitch estimation, within-domain effects

Further analyses were conducted to ensure that differences in cross-dimensional interference were not due to differences in the accuracy with which participants reproduced pitches. For each participant we computed the slope of the effect of the actual pitches on participants' reproduced pitches (Dutch Height: Slope=1.02, $t(19)=12.50$, $p=.0001$; Dutch Thickness: Slope=0.87, $t(19)=12.68$, $p=.0001$; Farsi Height: Slope=0.61, $t(19)=6.67$, $p=.0001$; Farsi Thickness: Slope=0.58, $t(19)=6.21$, $p=.0001$). According to a 2 x 2 ANOVA, Language did not interact with Task to predict the effect of actual pitch on estimated pitch, $F(1,79)=0.46$, $MSE=.002$, *ns*. Overall, Dutch speakers' pitch estimates were more accurate than Farsi speakers', $F(1,79)=16.96$, $MSE=.002$, $p=.0001$, but this main effect of Language on within-domain performance cannot explain the Language x Task interaction we found in the cross-domain analysis.

Finally, we conducted a 3-way ANOVA combining the cross-domain and within-domain analyses, after normalizing all values of space and pitch for each participant. There was a 3-way interaction of Language (Dutch, Farsi), Task (Height Interference, Thickness Interference) and Domain (Within-domain effects, Cross-domain effects); $F(1,159)=7.04$, $MSE=.0001$, $p=.01$, indicating that the predicted cross-dimensional interference effects cannot be explained by unpredicted differences in within-domain performance.

In summary, results suggest that people who use different metaphors in their native languages form correspondingly different mental representations of musical pitch.

Experiment 2: Eliminating Verbal Labeling

In Experiment 1, all stimuli and responses were nonlinguistic, and no language production or comprehension was required to perceive or reproduce the stimuli. Did participants nevertheless use language covertly to label the pitches? This is unlikely to account for the observed pattern of cross-dimensional

interference, for a combination of reasons. First, the increments of space and pitch were too fine-grained to be labeled helpfully using ordinary non-technical words: Labels such as 'high' and 'low' are too coarse to help participants discriminate nine randomly ordered pitches at the level of the semitone (e.g., the word 'low' applies to all of the pitches on the low end of the continuum, and therefore cannot help participants to discriminate among them). Second, and most importantly, covertly labeling pitches as 'high' or 'low' would work *against* the observed spatial interference effects, since space and pitch varied orthogonally.

Still, in order to rule out the possibility of online language effects experimentally, we conducted a version of the Height Interference task in Dutch speakers with the addition of a verbal interference condition. If the effect of height on pitch is driven by covertly labeling the stimuli using spatial words, then it should disappear under verbal interference. However, we hypothesized that this effect was not due to online use of spatial metaphors for pitch in language, but rather to the activation of an implicit association between nonlinguistic, analog representations of space and pitch in memory: a mental metaphor (Casasanto, 2010; Lakoff & Johnson, 1980). If so, the effect of height on pitch should persist under verbal interference.

Methods

Participants.

A new sample of native Dutch speakers (N=22) participated for payment.

Materials.

The same materials as in the Height Interference task were used. Additionally, 81 unpronounceable 5-letter strings were constructed.

Procedure.

After 8 practice trials, participants were presented with 81 line-pitch pairings, one at a time, and they sang back the pitches that they heard in the presence of task-irrelevant visuo-spatial information, as in Experiment 1. Additionally, before each trial they saw one of the 5-letter strings for 2 seconds,

and were instructed to rehearse the letters silently. After one third of the trials, participants' recognition of the letter strings was tested using a 2-alternative forced choice. Participants responded by pressing the key corresponding to either the correct string or a foil (S or L key). The next trial began after a 500 ms inter-trial interval. After 40 trials, participants had a self-paced break. Testing lasted about 20 minutes.

Results and Discussion

Verbal interference task.

Participants' recognition of the correct letter series was much greater than chance (Mean accuracy=85%, SD=9.25; $t(21)=17.76$, $p=.0001$), indicating that they were engaged in the verbal interference task.

Pitch reproduction task.

For each participant we computed the slope of the effect of normalized spatial height on reproduced pitch (Figure 1b). Spatial height still influenced pitch estimates, even under verbal interference, Slope =3.66, $t(21)=2.65$, $p=.02$ (Figure 1b, Figure 3). Slopes for the Dutch participants performing the Height Interference task in Experiment 1 did not differ from the slopes of participants performing the same task under verbal interference in Experiment 2 (difference of normalized slopes=1.01, $t(40)=.58$, ns). Effects of space on pitch cannot be attributed to covert activation of verbal labels for the stimuli.³

³ Although it is generally assumed that language is not playing an online role if a task is shown to be unaffected by concurrent verbal interference, we acknowledge that alternative interpretations of verbal interference exist (e.g., Lupyan, 2012).

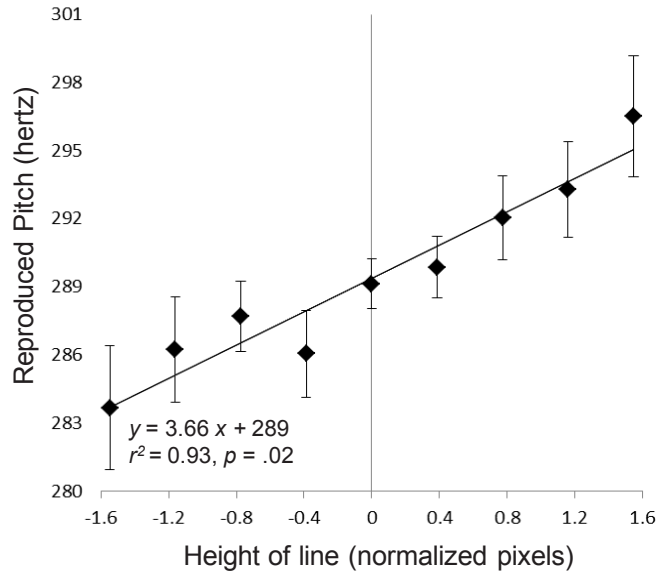


Figure 3. Results of Experiment 2. Spatial height still influenced pitch estimates, even under verbal interference. Error bars represent standard error of the mean.

Experiment 3: Does Language Shape Pitch Representations?

Although the data from Experiments 1 and 2 closely follow predictions based on linguistic metaphors, they are nevertheless correlational. A 2-part training study was conducted to investigate whether language can play a causal role in pitch representation.

First, Dutch speakers completed sentences about pitch relationships using Farsi-like thickness metaphors (Thickness Training), or using familiar height metaphors (Height Training) as a control. To determine whether this linguistic training influenced nonlinguistic pitch representations, we then tested all participants on the Thickness Interference task from Experiment 1. If experience using thickness-pitch metaphors in language causes Farsi speakers to rely on mental representations of spatial thickness to think about musical pitch, then repeatedly using similar linguistic metaphors during training should cause Dutch speakers to perform similarly to Farsi speakers on the nonlinguistic Thickness Interference task.

Methods

Participants.

Native Dutch speakers (N=60) participated for payment. Half were assigned to the Thickness Training task, and the other half to the Height Training task. The data of three additional Dutch-speaking participants could not be analyzed due to sound recording problems.

Materials and procedure.

During the training phase, participants completed 196 fill-in-the-blank sentences, typing the words *dunner* (thinner) and *dikker* (thicker) in the Thickness Training condition and the words *hoger* (higher) or *lager* (lower) in the Height Training condition. In both tasks, half of the sentences compared the spatial height or thickness of physical objects (e.g., A tower is higher / lower than a blade of grass; A pillar is thicker / thinner than a finger); the other half compared pitches of different sounds (e.g., A flute sounds higher / lower than a tuba; A flute sounds thicker / thinner than a tuba). Participants were not told whether 'thicker' or 'thinner' meant 'higher' or 'lower'; rather, they were left to infer the correct mapping based on 3 correctly-completed example sentences, and on the feedback they received after each trial, either 'goed' (correct) or 'fout' (incorrect). Training took about 20 minutes. After training phase, all participants performed the Thickness Interference task from Experiment 1.

Results and Discussion

Training Phase.

Participants filled in the blanks with high accuracy for both the Height Training (Mean accuracy=99%, SD=0.77) and the Thickness Training task (Mean accuracy=99%, SD=1.05). Accuracy did not differ between tasks (difference=0.2%, $t(58)=.96$, *ns*).

Test Phase.

Thickness influenced pitch reproduction following Thickness Training, Slope=1.45, $p=.003$, but not following Height Training, Slope=0.08, *ns*; difference of slopes=1.38, $t(58)=1.84$, $p=.07$ (Figure 4). The effect of thickness on pitch

reproduction in thickness-trained participants was statistically indistinguishable from the effect in native Farsi speakers, difference of slopes=1.39, $t(48)=1.12$, *ns*, and was greater than the effect in untrained Dutch speakers, difference of slopes=2.06, $t(48)=2.02$, $p=.05$.

Experience using space-pitch mappings in language can change non-linguistic mental representations of musical pitch. Using the ordinary space-pitch metaphors in ones' native language may shape pitch representations via learning mechanisms similar to those that changed people's representations in this laboratory training task.

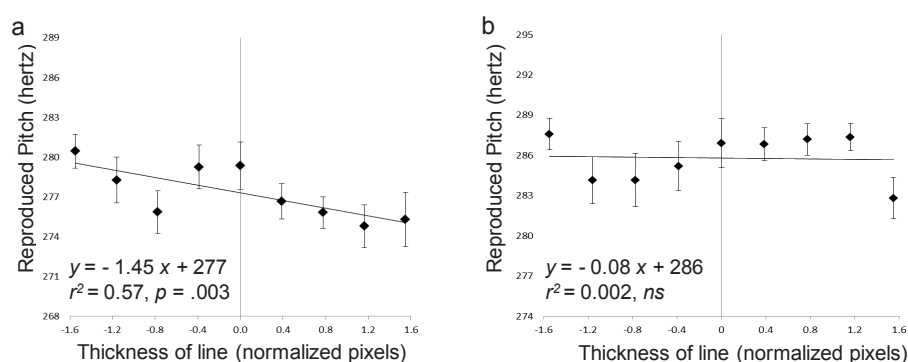


Figure 4. Results of Experiment 3. Effects of thickness interference on pitch estimates in speakers of Dutch after Thickness Training (left) and after Height Training (right). Thickness influenced pitch reproduction following Thickness Training, Slope=1.45, $p=.003$, but not following Height Training, Slope=0.08, *ns*. Error bars represent standard error of the mean.

Experiment 4: Does Language Create Space-Pitch Mappings?

What role does language play in shaping nonlinguistic pitch representations? On one proposal, perhaps language is instrumental in creating mental metaphors (Boroditsky, 2001; Gentner & Wolff, 2000). Using spatial words like 'high' or 'thick' metaphorically could encourage speakers to discover analogical correspondences between space and pitch that did not exist (or were not used) prior to exposure to linguistic metaphors.

Alternatively, using linguistic metaphors could modulate the strength of pre-existing mental metaphors (Casasanto, 2008). Consistent with this proposal, pre-linguistic infants are sensitive to the association between height and pitch found in English and Dutch metaphors (Walker, et al., 2010), and also to the association between thickness and pitch found in Zapotec and Farsi metaphors (Dolscheid, Hunnius, Casasanto, & Majid, 2012). The role of language, then, may be to adjust the relative strengths of pre-linguistic space-pitch associations. Suppose each time people produce or understand a space-pitch metaphor in language they activate the corresponding nonlinguistic association between space and pitch. Over time, speakers of a 'height' language like Dutch would strengthen the height-pitch mapping at the expense of the thickness-pitch mapping, due to competitive associative learning (Casasanto, 2008; 2010). The opposite would be true for speakers of a 'thickness' language like Farsi, whose linguistic experience would lead to a strengthening of the pre-existing thickness-pitch mapping at the expense of the pre-existing height-pitch mapping.

To distinguish these alternatives, in Experiment 4 we trained Dutch speakers to use a 'reverse Farsi' mapping, following the same procedure as in the Thickness Training condition of Experiment 3, with one exception: Rather than learning Farsi-like metaphors that associated high frequencies with 'thin' and low frequencies with 'thick', participants learned the opposite thickness-pitch mapping (i.e., low=thin, high=thick), which is not known to be conventionalized in any language.

If learning a new linguistic metaphor causes people to create a new space-pitch mapping, then training Dutch speakers to use the Reverse-Farsi mapping should be just as effective as training them to use the Farsi-like mapping, since the two mappings are equally novel and systematic. Alternatively, if learning a new linguistic metaphor influences pitch representations by strengthening a pre-existing space-pitch mapping, then training should be more effective when the new metaphor corresponds to one of the space-pitch mappings found in pre-linguistic infants than when it contradicts one of these mappings.

Methods

Participants.

Native Dutch speakers (N=30) participated for payment. The data of one additional Dutch-speaking participants could not be analyzed due to sound recording problems.

Materials and procedure.

The materials and procedure were identical to Experiment 3, with one exception: Participants were trained to use the reverse mapping between thickness and pitch in language (i.e., low=thin, high=thick), prior to performing the nonlinguistic Thickness Interference task from Experiment 1.

Results and Discussion

Training Phase.

Participants filled in the blanks with high accuracy for the Reversed Thickness Training (Mean accuracy=95%, SD=6.69). Comparing accuracy between Reversed Thickness and Thickness Training, however, showed that participants made significantly more errors during Reversed Thickness Training (difference=3.83%, $t(58) = 3.04$, $p=.01$). This indicates that learning the reversed thickness metaphor was more difficult than learning the Farsi-like thickness metaphor.

Test Phase.

Thickness did not influence pitch reproduction following Reversed Thickness Training, (Slope=.34, *ns*; Figure 5). Comparing effects between Reversed Thickness Training and Thickness Training, the effect of thickness on pitch reproduction in thickness-trained participants was significantly greater than the effect in reversed-trained participants, difference of slopes=1.12, $t(58)=2.92$, $p=.03$.

Results of the Reversed Thickness Training suggest that language did not create a Farsi-like space-pitch mapping in Experiment 3. Rather, using Farsi-like linguistic metaphors strengthened the 'low=thick, high=thin' mapping that was

not evident in adult Dutch speakers' language or thought (see Experiment 1), but which has been observed in pre-linguistic infants (Dolscheid, et al., 2012).

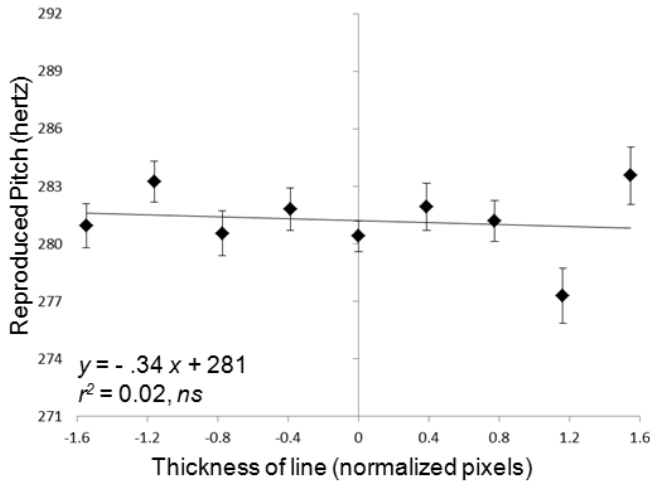


Figure 5. Results of Experiment 4. Effects of thickness interference on pitch estimates in speakers of Dutch after Reversed Thickness Training. Thickness did not influence pitch reproduction following Reversed Thickness Training, (Slope=.34, *ns*). Error bars represent standard error of the mean.

General Discussion

Dutch and Farsi speakers, who use different metaphors for pitch in language, also form correspondingly different nonlinguistic pitch representations. We show this via a double-dissociation between Dutch- and Farsi-speakers' performance on a pair of nonlinguistic psychophysical tasks. Dutch speakers, who talk about pitches as 'high' and 'low', incorporated irrelevant height information into their pitch estimates (but ignored irrelevant thickness information). Farsi speakers, who talk about pitches as 'thin' and 'thick', incorporated irrelevant thickness information into their pitch estimates (but ignored irrelevant height information). When Dutch speakers were trained to use Farsi-like metaphors, they showed the same pattern of cross-dimensional thickness interference as native Farsi speakers. Beyond demonstrating a

language-thought correlation, results show that metaphors in language can play a causal role in shaping nonlinguistic mental representations of musical pitch.

The influence of spatial height on pitch estimates was found even when Dutch-speaking participants performed a concurrent verbal interference task, suggesting that effects of space on pitch were not mediated by the covert use of language during the psychophysical tasks. Rather, we propose that the effects of language on pitch representation occurred prior to testing: Using verbal 'height' or 'thickness' metaphors strengthened either the height-pitch or the thickness-pitch mapping in participants' memories, consequently weakening the alternative mapping due to some form of competitive associative learning (Casasanto, 2008).

During ordinary language use the relative strengths of height-pitch and thickness-pitch mappings may be modulated slowly, as instances of space-pitch metaphors accumulate over time. During the laboratory training task (Experiment 3), however, the relative strengths of these mappings were modulated quickly -- and we assume transiently -- because participants receive a concentrated 'dose' of the relevant linguistic metaphor, probably equivalent to weeks or months of normal language use. Presumably a thickness-pitch mapping was found in thickness-trained Dutch speakers because, at the time of test, they had experienced very frequent, very recent activation of the thickness-pitch mapping. (For compatible evidence that mental metaphors can be rapidly retrained, see Boroditsky, 2001; Casasanto, 2008; Casasanto & Bottini, 2010; Casasanto & Chrysikou, 2011; Fischer, Mills, & Shaki, 2010).

Beyond 'thinking for speaking'

On one influential view of the relationship between language and thought, patterns in language can influence nonlinguistic mental representations only (or primarily) while people are packaging their thoughts into words (Slobin, 1996), or while they are performing tasks for which verbal codes can be helpful (e.g., Gennari, Sloman, Malt, & Fitch, 2002; Papafragou, Hulbert, & Trueswell, 2008). But these online effects of language on high-level language-mediated thinking are only one sort of linguistic relativity effect. The present results support the proposal that language can also influence people's low-level perceptuo-motor

abilities, such as their ability to reproduce musical pitches, and that cross-linguistic differences in mental representation can be observed even when people are not using language online, overtly or covertly. This effect contrasts with other 'Whorfian' effects reported previously, for example in the domain of color, which disappear under verbal interference (e.g., Winawer et al., 2007).

Origins of space-pitch mappings

Pre-linguistic infants appear to be sensitive to both height-pitch and thickness-pitch mappings (Dolscheid, et al., 2012; Walker, et al., 2010), suggesting that language does not create these space-pitch mappings. Rather, we propose that using verbal metaphors strengthens one of these preexisting mental metaphors (while weakening the other). Consistent with this proposal, training adults to use new verbal metaphors was more effective in changing their nonlinguistic pitch representations when the new metaphors corresponded to pre-verbal space-pitch mappings than when they contradicted these mappings. Dutch speakers quickly learned and used the thickness-pitch mapping found in Farsi (Experiment 3), but not the reverse mapping (Experiment 4), which is not known to be encoded in any language.

Where do these space-pitch mappings come from, if not from language? On one possibility, both the height-pitch and thickness-pitch mappings could reflect innate cross-modal correspondences, with no experiential basis (Walker, et al., 2010). Alternatively, both of these space-pitch mappings could plausibly be based on correspondences in the physical world. The relationship between thickness and pitch is evident in musical instruments (e.g., thicker strings produce lower tones; Shayan, et al., 2011). The relationship between height and pitch is evident in bodily experience: As people produce higher pitches the larynx rises, and as they produce lower pitches it descends (Miller, 1986).

Yet, just-so stories about the physical origins of mental metaphors should be interpreted with caution. It is easy to find other physical regularities that predict different relationships between pitch and space (e.g., taller people tend to have *lower* voices). It remains an open question to what extent space-pitch mappings in our minds emerge over developmental time, as individuals track experiential regularities (Lakoff & Johnson, 1999), or over evolutionary time as

the neural substrates of spatial cognition were exapted for non-spatial functions (Pinker, 1997).

Conclusions

Whatever their origins may be, different space-pitch mappings have become encoded in the languages we speak, in expressions that are so highly conventionalized speakers may hardly notice they are using spatial metaphors. Yet, language-specific metaphors shape people's nonlinguistic representations of musical pitch. As a result of habitually using one spatial metaphor or another, speakers of different languages tend to form systematically different representations of the same physical experiences, even when they are not using language.

Acknowledgements

We like to thank Ozge Ozturk for discussions, Francisco Torreira for help with Praat analyses and Tomas Bergvelt for translations.

The sound of thickness: Prelinguistic infants' associations of space and pitch

Chapter 3

Based on:

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.), *CogSci 2012: Proceedings of the 34th Annual Meeting of the Cognitive Science Society* (pp. 306-311). Austin, TX: Cognitive Science Society.

Abstract

People often talk about musical pitch in terms of spatial metaphors. In English, for instance, pitches can be 'high' or 'low', whereas in other languages pitches are described as 'thick' or 'thin'. According to psychophysical studies, metaphors in language can also shape people's nonlinguistic space-pitch representations. But does language establish mappings between space and pitch in the first place or does it modify preexisting associations? Here we tested 4-month-old Dutch infants' sensitivity to height-pitch and thickness-pitch mappings in two preferential looking tasks. Dutch infants looked significantly longer at cross-modally congruent stimuli in both experiments, indicating that infants are sensitive to space-pitch associations prior to language. This early presence of space-pitch mappings suggests that these associations do not originate from language. Rather, language may build upon pre-existing mappings and change them gradually via some form of competitive associative learning.

Introduction

Does a cake taste yellow? Or a tone played by a trumpet sound scarlet? For some people they do. Yet synesthesia, a condition in which stimulation of one sensory modality induces systematic perceptual experiences in another modality, is relatively rare. Other types of cross-modal associations, however, can be found in non-synesthetes too. Psychophysical studies have shown that adults and children without synesthesia associate higher pitches with sharper edges (Marks, 1987; Parise & Spence, 2009), positions higher in space (e.g., Ben-Artzi & Marks, 1995; Melara & O'Brien, 1987), lighter color (Hubbard 1996; Marks 1989; Melara 1989), and increasing brightness (Marks, 1987).

Even infants seem to be sensitive to some of these associations. Cross-modal correspondences between loudness and brightness have been demonstrated in 20- to 30-day-old infants (Lewkowicz & Turkewitz, 1980). In a preferential looking task, 3- to 4-month-old infants preferred congruent trials – in which visuospatial height and pitch height corresponded – over incongruent trials (Walker et al., 2010). Infants looked longer at a ball moving upwards if it was accompanied by a rising pitch than if it was accompanied by a falling pitch. Also, prelinguistic infants under 1-year-old 'matched' visual arrows pointing up or down with tones sweeping up or down in frequency (Wagner, Winner, Cicchetti, & Gardner, 1981).

These findings have led to the assumption that cross-modal mappings are innately hardwired in the brain (Mondloch & Maurer, 2004) and represent an unlearned aspect of perception (Walker et al., 2010). Accordingly, some of these associations are posited to be universal (Marks, Hammeal, & Bornstein, 1987; see also Spence, 2011).

However, there are other findings that seem to be at odds with these conclusions. A number of cross-modal correspondences are only acquired later in the course of development. For instance, even 9-year-old children are not able to systematically match size and pitch, a task that is consistently solved by adults (Marks et al., 1987; but see Mondloch & Maurer, 2004). Cross-modal correspondences are also affected by developmental changes. Unlike adults and older children, 2-year-olds consistently map light grey to smaller objects and dark grey to bigger objects (Smith & Sera, 1992, but also see Walker & Walker,

2012). In the course of language acquisition, however, children's associations gradually shift. As a result, it has been suggested that the trajectory of cross-modal relations may be affected by language (Smith & Sera, 1992).

Language, on one hand, appears to mirror cross-modal experience. The auditory domain, for instance, is often linguistically encoded in terms of other sensory modalities (Williams, 1976). People use metaphors like 'soft' voice, 'dark' timbre or 'high' pitch, suggesting that language echoes cross-modal perceptual impressions. On the other hand, language also seems to affect cross-modal associations. For example, Martino and Marks (1999) propose that cross-modal effects, like the association between space and pitch, may be mediated by language. Various tasks show correspondences between spatial height and pitch consistent with high-low metaphors in language (Rusconi, et al, 2006, Evans & Treisman, 2010). However, since linguistic labels and height-pitch associations merely coincide, the direction of influence is hard to establish and the contribution of language remains unclear.

Cross-linguistic comparison provides one way to overcome this limitation. Not every language uses the same metaphors for pitch (Ashley, 2004; Levinson & Majid, 2007). For example, while languages like English and Dutch talk about pitch in terms of 'height', other languages like, Farsi, Turkish and Zapotec (spoken in Mexico) describe high-frequency pitch as 'thin' and low-frequency pitches as 'thick' (Shayan, Ozturk, & Sicoli, 2011). To find out whether these differences in spoken metaphors correspond to different mental representations of pitch, Dolscheid and colleagues, conducted a series of nonlinguistic psychophysical experiments in adult speakers of Dutch (a 'height' language) and Farsi (a 'thickness' language). Participants were asked to reproduce musical pitches that they heard in the presence of irrelevant spatial information, (i.e. lines varying either in height or in thickness) (Dolscheid, Shayan, Majid, & Casasanto, 2013, see Chapter 2). Dutch speakers' pitch estimates were significantly modulated by spatial height but not by thickness. Conversely, Farsi speakers' pitch estimates were modulated by spatial thickness but not by height. Overall, the results indicated that nonlinguistic pitch-space associations follow language-specific vocabulary, suggesting that cross-modal pitch representations can be affected by language.

At this point, however, it is unclear whether language establishes cross-modal mappings between space and pitch in the first place, or whether it merely modifies preexisting associations. While some researchers stress the relevance of language in concept formation (e.g., Gopnik & Meltzoff, 1997), others argue that conceptual representations must precede the acquisition of language (e.g., Bloom, 2000; Bloom & Keil, 2001). The former position allows for a stronger role of language in space-pitch associations; while the latter position suggests that children are likely to have some notion of space-pitch correspondences prior to learning language. Consistent with this latter view, infants seem to be sensitive to height-pitch mappings even prelinguistically (Walker et al., 2010). Critically, however, nothing is known about the origins of thickness-pitch relationships. Do children also have this mapping available to them prior to language, or is it only learned on the basis of language-input?

One possibility is that children could start out with both a height-pitch and a thickness-pitch mapping even before they learn language. The strength of these mappings might then subsequently be adjusted, according to the relative frequencies of space-pitch metaphors in the languages children acquire (Casasanto 2008, 2010; Dolscheid et al., 2013; Smith & Sera, 1992). Alternatively, height-pitch and thickness-pitch associations might follow different trajectories. Whereas height-pitch associations are available to prelinguistic infants, the thickness-pitch mapping might only be learned later. Metaphors in language could provide one possible way to learn this association. Using thickness terminology to refer to pitch may invite speakers to align correspondent representations and extract similarities between space and pitch in a process called structural alignment (see e.g., Boroditsky, 2001; Gentner, 2003). In line with this proposal, Shayan et al. (submitted) found that Turkish and Farsi 2- to 5-year-olds were able to successfully map thickness to pitch but same-aged German children (who like English and Dutch speakers do not have a thickness metaphor) were not able to make this association successfully. This is consistent with the proposal that language input promotes cross-modal associations between thickness and pitch. Note, however, that these results do not rule out the possibility that the thickness-pitch mappings were available to all infants prelinguistically, but are no longer equally available to German children.

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

In a second step, we tested prelinguistic infants in a thickness-pitch task analogous to the height-pitch task. Instead of balls moving up and down the screen, a vertical tube varied in thickness, changing continuously from thin to thick. In the congruent condition, the pitch of the sound 'rose' and 'fell' in accordance with the movement of the tube (e.g., the tube expanded when the pitch 'fell'). In the incongruent condition the pitch of the sound 'rose' and 'fell' in opposition to the movement of the tube (e.g., the tube expanded when the pitch 'rose') (see Fig. 1b).

We reasoned that if both, height-pitch and thickness-pitch mappings are available to infants prelinguistically, infants should prefer both congruent height-pitch and congruent thickness-pitch stimuli over incongruent ones. If however, height-pitch and thickness-pitch relationships follow different developmental trajectories, with thickness mappings only becoming acquired later, then infants should show preferences for congruent height-pitch stimuli but not for congruent thickness-pitch stimuli.

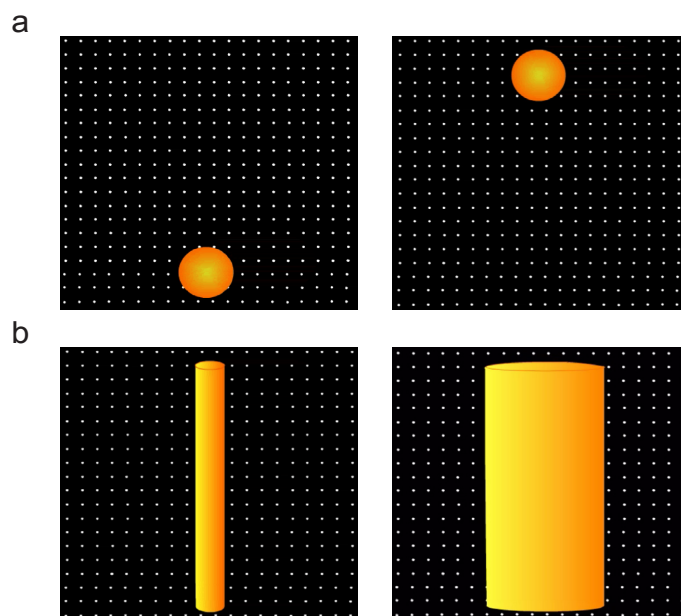


Figure 1: Examples of animations presented as stimuli in Experiment 1 (Panel a) and Experiment 2 (Panel b). In (a) the extremes of the ball's vertical trajectory are shown. In (b) the extremes of thickness are depicted. The images are reproduced to scale.

Experiment 1: Auditory pitch and visuospatial height

Methods

Participants.

Ten male and ten female infants completed the first (pitch-height) experiment (mean age = 129 days, range: 113 to 138 days). Another eight infants were tested, but not included in the analyses: two infants were excluded due to experimenter error; a further six infants were excluded due to fussiness.

Materials and Procedure.

QuickTime animations were presented on a 102 x 76 cm Sony LCD screen using HABIT software. Animations appeared within a 67 x 67 cm screen area ($25.6^\circ \times 25.6^\circ$), and lasted a maximum of 60 s. Before each animation, a flashing light appeared to ensure that infants attended to the screen. Infants sat in a Maxi-Cosi infant seat which was placed on their parent's lap, viewing the animations from a distance of 1.50 m. Infants' visual fixations were monitored and recorded

on video. Animations were stopped if the infant looked elsewhere for a single period of 1 s or more. The total duration an infant looked directly at the animation was determined by a subsequent frame-by-frame coding of the digitized video using SuperCoder. Coding was performed by a coder blind to the experimental condition. 25 percent of the data was double-coded by a second person, also blind to the condition.

Infants watched a 10-cm (4°) diameter orange ball moving up and down a 50-cm vertical trajectory in front of a 20 x 20 grid of small, white dots on a black field. The ball moved at a constant speed of 20 cm/s and paused for 42 ms at each endpoint. Animations were accompanied by the sound of a sliding whistle (a sinusoidal tone). The fundamental frequency of the sound changed at a constant rate, between 300 and 1700 Hz over 2.5 s, coinciding with a single phase of the animation (e.g., the ball moving up). The amplitude of the sound increased and then decreased between 47 and 84 dB within each phase of the animation, peaking at 1000 Hz. Amplitude thus changed about twice as fast as pitch to ensure that variation in perceived pitch and loudness were not confounded.

Every infant viewed three congruent animations interleaved with three incongruent animations. Half of the children watched a congruent animation first and the other half watched an incongruent animation first. During the whole experiment, parents were listening to music via headphones. Since parents could not hear the sliding sounds, they were unable to distinguish between experimental conditions (the spatial trajectories of the stimuli did not differ between conditions). We therefore ensured parents could not bias their infant's looking behavior.

Results

A high level of agreement was confirmed between the two observers in their coding of each infant's individual looking times (Pearson's $r(28) = .99$, $p < .001$).

14 of the 20 infants looked longer at the congruent animation than at the incongruent animation. On average, infants looked at the congruent animations for 31.7 s ($SD = 11.4$) and at the incongruent animations for 26.1 s ($SD = 13.3$). A

paired-samples t-test confirmed that infants looked significantly longer at the congruent animations, $t(19) = 1.99$, $p = .03$, $d = 0.45$ (one-tailed) (see Figure 2, panel a).

Experiment 2: Auditory pitch and visuospatial thickness

Methods

Participants.

Ten male and ten female infants completed the second (pitch-thickness) experiment (mean age = 127 days, range: 113 to 138 days). An additional eight infants were tested, but their data was not analyzed: one infant was excluded due to technical problems; a further 7 infants were excluded due to fussiness.

Materials and Procedure.

The same procedure as in Experiment 1 was used. This time, infants watched a vertical orange tube that varied in thickness, changing continuously from thin to thick (see Figure 1).

The animation was presented on a 20 x 20 grid of small, white dots on a black field, as in Experiment 1. The tube was 60 cm long ranging from 6 to 26 cm in width. It expanded at a constant speed of 8 cm/s and paused for 42 ms at each endpoint. Animations were accompanied by the sound of the exact same sliding whistle as in Experiment 1. The fundamental frequency of the sound changed at a constant rate, between 300 and 1700 Hz over 2.5 s, coinciding with a single phase of the animation (i.e. during tube expansion). Each infant viewed three congruent animations interleaved with three incongruent animations, with half of the children watching a congruent animation first and the other half watching an incongruent animation first. During the whole experiment, parents were listening to music via headphones.

Results

A high level of agreement was confirmed between the two judges in their estimates of each infant's individual looking times (Pearson's $r(28) = .99$, $p < .001$).

13 of the 20 infants looked longer at the congruent animations than the incongruent animations. On average, infants looked longer at the congruent 24.4 s ($SD = 11.8$) than the incongruent animations 19.4 s ($SD = 11.5$). A paired-samples t-test confirmed that infants looked significantly longer at the congruent animations, $t(19) = 2.19$, $p = .02$, $d = 0.43$ (one-tailed) (see Figure 2, panel b).

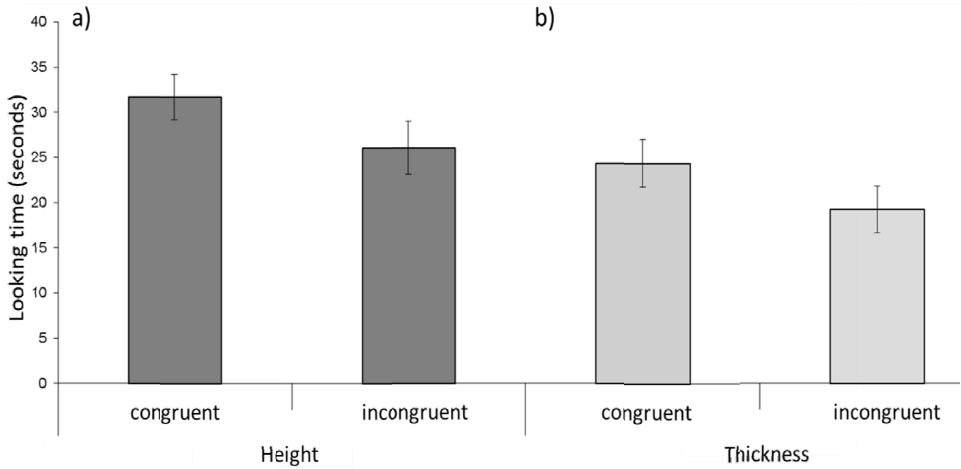


Figure 2: Results of Experiment 1 (Panel a) and Experiment 2 (Panel b). Infants looked significantly longer at congruent stimuli in both experiments.

Between experiment comparison

Dutch infants looked significantly longer at cross-modally congruent stimuli in both experiments. While this suggests a comparable starting point for both thickness-pitch and height-pitch mappings, it is nevertheless possible that infants display differential preference with respect to the two mappings. We therefore compared the results of the two previous experiments directly.

Results

Submitting looking times to a 2 (Space: height vs. thickness) by 2 (Congruency: congruent vs. incongruent) mixed ANOVA yielded a significant main effect of Space ($F(2,38) = 4.40$, $p = .04$, $\eta_p^2 = 0.10$), showing that looking

times differed between height and thickness stimuli. Infants looked longer at height stimuli as compared to thickness stimuli, indicating that perhaps height was more salient for them. However, no interaction between Space and Congruency ($F(2,38) = 0.03$, ns , $\eta_p^2 = 0.00$) was observed. There was thus no indication that looking time congruency effects differed between the two experiments. In line with this, percentage reduction in looking time across experiments was of comparable size, i.e., 18% for the height-pitch experiment and 20% for the thickness-pitch experiment.

General Discussion

Our results demonstrate that prelinguistic infants are sensitive to correspondences between auditory pitch and spatial information of two different types, visuospatial thickness as well as height. Dutch infants looked significantly longer at cross-modally congruent stimuli in both experiments, suggesting a comparable starting point for height-pitch mappings and thickness-pitch mappings.

It is possible that these mappings are only present in very young infants but get lost in the course of development due to neuronal pruning. Whereas 2-to 3-month-old infants were found to be sensitive to arbitrary associations between colors and shapes, 8-month-old infants no longer show this early synesthetic association (Wagner & Dobkins, 2011). Does the same developmental trajectory hold true for space-pitch mappings? Unlike synesthetic color-shape associations that seem highly individualized and thus unspecific (e.g., one infant might associate triangles with green, and another with red), space-pitch associations follow a specific pattern, showing the same congruity preferences found in languages. It is therefore possible that space-pitch mappings persist during infancy and childhood. In line with this suggestion, sensitivity to height-pitch associations has been reported in 6-month-olds (Braaten, 1993, as cited in Spence, 2011) as well as in children aged 4 to 5 years (Roffler & Butler, 1967). On the other hand, 2- to 5-year-old German speaking children have been found to be insensitive to the thickness mapping (Shayan et al., submitted). There is also contradictory evidence regarding children's sensitivity to size-pitch associations. Whereas Marks et al. (1987) report that children are unable to systematically

map size (big vs. small) to pitch until they are 13 years old, Mondloch and Maurer (2004) find evidence for size-pitch congruency effects in children as young as 3 years of age. In yet another study, 10-month-old infants were found to be insensitive to size-pitch correspondences (Haryu & Kajikawa, 2012).⁴ Details about the developmental trajectory of space-pitch mappings thus remain unclear and are subject to future research.

One aspect that seems to facilitate the detection of cross-modal associations is motion (see also Jeschonek, Pauen & Babocsai, 2012). Mondloch and Maurer presented children with moving balls that differed in size; while Marks et al. and Shayan et al. used static stimuli. In the present study, too, the dynamic display of spatial information (up- and downward movement or horizontal expansion) and pitch (presented as glides) may have facilitated the detection of corresponding information. Displaying stimuli dynamically and in synchronicity could direct infants' attention to the relational correspondences across modalities. However, movement by itself cannot explain the pattern of results: infants must still align stimuli attributes that are congruent to each other.

Language acquisition and cross-modal associations

Our findings demonstrate that both height-pitch and thickness-pitch correspondences are perceived before the infant has mastery of language. While this finding is consistent with the view that representations precede language (e.g., Bloom, 2000), it does not entail that these associations are fixed. Language could still influence the structure and content of preexisting mental representations via simple learning mechanisms. In the course of language acquisition, the relative strengths of different space-pitch mappings could be adjusted according to the language-specific frequencies of metaphors that

⁴ The thickness-stimuli used in Experiment 2 could also be interpreted as a size manipulation. Indeed, even though movement was restricted to the horizontal plane which is characteristic for thickness, there is a concomitant difference in overall size. For the present purposes, nothing rests on being able to make the distinction between thickness and size, per se, since the reported consistency in the ability to make the cross-modal mapping to pitch applies equally to both spatial parameters.

children acquire (Casasanto 2008, 2010). Over time, speakers of a 'height' language like Dutch would strengthen the height-pitch mapping at the expense of the thickness-pitch mapping – and vice versa for speakers of a 'thickness' language like Farsi (Dolscheid et al., 2013). Evidence in support of this associative learning account is provided by linguistic training experiments. Adult Dutch speakers, after being trained to use Farsi-like metaphors describing pitch relationships in terms of thickness, demonstrated nonlinguistic thickness-pitch mappings just like Farsi speakers. By contrast, when participants received the same amount of linguistic training with an alternative space-pitch mapping that is not present in any known language, they showed no effect of training (Dolscheid et al., 2013, see Chapter 2). These training studies demonstrate a causal role for language in strengthening the use of some nonlinguistic mappings more than others.

While language may enforce particular pitch-space mappings, this proposal has to take into account that metaphors pose additional demands in language acquisition. Pitch metaphors are inherently polysemous; the acquisition of both spatial and sound meanings is likely more complex than when a single meaning has to be acquired (see e.g., Johnson, 1992). Consistent with this proposal, French speaking children trained to describe sounds using either the single-meaning terms *aigu* and *grave* (a pair of antonyms used only to label high and low pitches) versus the polysemous words *haut* and *bas* (which are used to refer to pitch and space) were better able to label sound stimuli (Costa-Giomi & Descombes, 1996).

Aside from polysemy, another important attribute of metaphorical language lies in its directionality. Taking spatial metaphors of time as an example, people talk about time in terms of space far more often ('a long vacation'; 'a short meeting') than they talk about space in terms of time (though it occasionally occurs: 'I live two minutes from here') (Casasanto, 2008, 2010; Lakoff & Johnson, 1980). For pitch metaphors the same asymmetry seems to hold, which has also been found to be reflected in adults' nonlinguistic pitch representations (Casasanto, 2010). Note, however, that our results are agnostic of a space-pitch asymmetry in prelinguistic infants. While we have demonstrated that infants are able to detect space-pitch associations, our tasks do not speak to

possible directionality. Future studies are necessary to determine whether language plays a role in introducing this asymmetry (e.g., see Merrit, Casasanto, & Brannon, 2010), or whether it is present independent of language (e.g., see Marks et al., 1987).

Effects of cross-modal associations on language

Since cross-modal associations are present before children acquire language, it is possible that the associations themselves shape metaphors in language. We find the height-pitch metaphor in languages such as Spanish, German and Polish (Rusconi et al., 2006), as well as non-Indo-European languages, like Japanese and Chinese. In all of these languages, 'high' refers to high frequency sounds and 'low' to low frequency sounds, but not the reverse (however, examples of pitch vocabulary with a contrasting mapping have been reported in 'Are'are, see Zemp & Malkus, 1979). Likewise, psychophysical studies demonstrated that participants associate higher pitches with smaller objects, not with larger objects (e.g., Gallace & Spence, 2006). For the Kpelle and Jabo people in Liberia, this association is also encoded in language: 'small' refers to high pitch and 'big' refers to low pitch (see e.g., Eitan & Timmers, 2010). Prelinguistic associations, alongside correlations of properties in the real world, may thus serve as guiding principles that constrain the way pitch gets lexicalized, across languages.

Consequently, it might be harder to learn linguistic metaphors that are inconsistent with cross-modal mappings for which there is evidence in the natural world. The results of a training study support this suggestion (see Chapter 2). Dutch speakers trained to use reversed thickness-pitch mappings (thick=high, thin=low) were not able to master this association, even though they could learn the comparable congruent mapping (Dolscheid et al., 2013). It thus appears that language cannot easily retrain mappings that are supported by correlations present prelinguistically and/or supported by real world experience. Early sensitivity to certain mappings might therefore constrain the set of cross-modal associations that are likely to be observed in language and the mind.

Origins of cross-modal mappings?

Are cross-modal mappings innate? Based on the current evidence, we can only conclude that cross-modal associations between space and pitch are present from very early. By the age of 4 months, however, infants may well have encountered enough relevant co-occurrences in their interaction with the world to have learned these mappings. Thickness-pitch mappings seem especially prevalent: thicker strings produce lower notes, bigger bells have lower chimes, and people with bigger ('thicker') bodies tend to have lower voices. While infants may have internalized these regularities, the case for height-pitch mappings (that do not necessarily follow correspondences in the real world) is not conclusive (see also Walker et al., 2010).

Conclusions

No matter whether cross-modal associations are inborn or learned, the finding that both height-pitch and thickness-pitch mappings can be observed in infants as young as 4 months of age constrains theorizing about the role that language plays in shaping nonlinguistic mental representations of pitch.

Our data show that space-pitch associations are present prior to language, suggesting that language is unlikely to create cross-modal mappings between space and pitch, even if language seems to play this role in other domains (Gentner, 2002).

It appears that both the height-pitch mapping found in languages like Dutch and the thickness-pitch mapping found in languages like Farsi are already present in prelinguistic infants' minds. This finding suggests that people who use different spatial metaphors for pitch in their native languages come to think about pitch differently not because language instills in them one cross-modal mapping instead of the other, but rather because language strengthens one pre-existing mapping at the expense the other, via some form of competitive associative learning (Casasanto, 2008, 2010; Dolscheid, et al., 2013). The precise learning mechanisms that give rise to cross-linguistic differences in pitch representation, and the underlying neural mechanisms, remain topics for future research.

Acknowledgments

We would like to thank Sho Tsuji, Manu Schuetze, Marjolijn van Gelder, Angela Khadar, Margret van Beuningen, Nienke Dijkstra, Laura Arendsen, Dirkje van der Aa and Webb Phillips.

Space and pitch in language and thought: Exploring the locus of metaphor

Chapter 4

Based on:

Dolscheid, S., & Casasanto, D. (in preparation). Space and pitch in language and thought: Exploring the locus of metaphor

Abstract

According to conceptual metaphor theory, metaphorical mappings are not 'in' language. Rather, linguistic metaphors like 'a *high* pitch' reflect the activation of a particular non-linguistic spatial schema in a speaker's mind, and cue the activation of this schema in a listener's mind. This proposal makes a counterintuitive prediction: It should be possible to activate mental metaphors using the *wrong words*, so long as the words activate the *right schema*. To test this prediction, we asked English speakers to perform a space-pitch congruity task in which they classified high- and low-frequency pitches using the conventional terms 'high' and 'low' or using the words 'tall' and 'short', which also refer to the upper and lower poles of a vertical spatial continuum. Participants were faster to classify high pitches as 'high' and low pitches as 'low' (congruent mapping) than to classify high pitches as 'low' and low pitches as 'high' (incongruent mapping), consistent with previous findings. A similar congruity effect was found when participants classified pitches as 'tall' and 'short', even though these words cannot be used sensibly in English to refer to pitch (e.g., a *tall* soprano does not mean a *high* soprano). A further experiment showed that this effect was specific to spatial words that cued a vertical spatial schema, and that it was not an artifact of linguistic markedness or polarity alignment. Lexically inappropriate words can activate mental metaphors for pitch so long as they cue the right spatial schema, suggesting that the locus of space-pitch metaphors is in thought rather than in language.

Introduction

"In short, the locus of metaphor is not in language at all, but in the way we conceptualize one mental domain in terms of another." (Lakoff, 1993, p. 203).

This central claim of conceptual metaphor theory entails that people not only use metaphors in language but that they also *think* in terms of mental metaphors. Whereas support in favor of this proposal was initially linguistic (Lakoff & Johnson, 1980) and thus circular in nature (McGlone, 2007), there is now abundant experimental evidence demonstrating metaphorical thinking in domains like preference (Casasanto, 2009), social dominance (Schubert, 2005), and time (Boroditsky, 2000). However, in most experimental tests of metaphor theory, participants are still required (or at least have the opportunity) to process stimuli in language.

Dolscheid, Shayan, Majid, and Casasanto (2013, see Chapter 2) provide the first evidence that people think metaphorically even when they are prevented from using linguistic metaphors. Dutch speakers who conventionally talk about musical pitch in terms of spatial height also activated height-pitch mappings during a completely non-linguistic psychophysical task. Participants were asked to reproduce musical pitches that they heard in the presence of spatial information (i.e. lines that varied in spatial height). Even though spatial variation was orthogonal to variation in pitch and was task-irrelevant, participants' pitch estimates were significantly modulated by spatial height (the higher the line, the higher the reproduced pitch on average). The task was designed such that the predicted pattern could not result from covert use of space-pitch metaphors in language: Covertly labeling pitches as 'high' or 'low' would have worked against the observed spatial interference effects. Moreover, the effect of space-pitch interference persisted even when pitches were heard and reproduced with concurrent verbal interference (Dolscheid et al., 2013, see Chapter 2).

Implicit height-pitch mappings have been found in prelinguistic infants (Chapter 3). Three- to 4-month-old infants have been found to look longer at congruent trials (in which visuospatial height and pitch height were aligned) compared to incongruent trials in a preferential looking task (Walker et al., 2010;

Dolscheid, Hunnius, Casasanto, & Majid, 2012). Taken together, the prelinguistic sensitivity to height-pitch mappings and their persistence in entirely nonlinguistic psychophysical experiments indicate that height-pitch associations can be activated independently of language.

However, there is an apparent tension between this conclusion and the conclusion that follows from Dolscheid et al.'s further work, which shows that people who use different space-pitch metaphors in their native languages also mentally represent pitch differently, according to the ways they speak (Dolscheid et al., 2013, Chapter 2). Whereas Dutch speakers talk about pitch in terms of spatial height, Farsi speakers describe high-frequency tones as being 'thin' and low-frequency tones as 'thick'. To find out whether cross-linguistic differences were reflected in thought, Dutch and Farsi speakers were tested on two nonlinguistic psychophysical space-pitch interference tasks. They were asked to reproduce pitches that they heard in the presence of lines that varied in spatial height, as described above (height-interference task) or in the presence of lines that varied in spatial thickness (thickness interference task). Results showed language-specific patterns of interference effects. Whereas Dutch speakers were significantly influenced by spatial height but not by thickness, the reverse was true for Farsi speakers. Further training experiments showed that non-linguistic pitch representations not only correlated with linguistic metaphors, but that using thickness-pitch language could also change these representations (see Chapter 2). Yet, although thickness-pitch mappings can be influenced by metaphors in language, there is also evidence that thickness-pitch associations (like height-pitch associations) are already present in the minds of 4-month-old infants (Dolscheid et al., 2012; for height-pitch associations also see Walker et al., 2010). It is therefore unlikely that these mappings arise on the basis of linguistic experience.

But if metaphors are not 'in language', and if they can arise prelinguistically, then why do people who use different linguistic metaphors also use correspondingly different mental metaphors? Dolscheid et al. (2013) propose the following answer regarding space-pitch metaphors in language and thought (see Casasanto 2008 for a similar proposal regarding space-time metaphors). Each time people produce or understand a space-pitch metaphor in language

they activate the corresponding nonlinguistic association between space and pitch (a mental metaphor). Over time, speakers of a 'height language' like Dutch strengthen the height-pitch mapping at the expense of the thickness-pitch mapping (and vice versa for speakers of a 'thickness language' like Farsi), due to some form of associative learning in which mappings that get practiced are strengthened at the expense of mappings that do not get practiced to the same extent. Thus, in the case of space and pitch, metaphorical language shapes metaphorical thinking because verbal expressions serve as cues to activate a particular spatial schema. For Dutch speakers, linguistic metaphors for pitch activate an analog spatial continuum with a particular dimensionality (i.e., 1-dimensional) and orientation (i.e., vertical), which maps to the pitch continuum with a particular polarity (i.e., higher-frequency pitches are higher in space, lower-frequency pitches are lower in space).

This proposal makes a prediction: If expressions like 'a high descant' and 'a low bass line' are cues to activate particular spatial representations (i.e., representations of positions along a vertical spatial continuum), and those representations are the active 'ingredient' in a mental metaphor, then any words that activate the same (or similar) spatial representations should also suffice to activate the familiar space-pitch mapping. That is, if the locus of metaphor is not in language, it should be possible to activate familiar mental metaphors by using the wrong words, so long as they cue the right spatial schemas.

Unlike high and low, tall and short cannot be used in conventional English to describe the height (i.e., frequency) of musical pitches. It is often the case that spatial metaphors in language can be freely extended, and novel coinages can be understood clearly and immediately. For example, a melodic line can be described not only as moving 'higher' or 'lower' but might also be described as rising, falling, swooping, dipping, soaring, or plummeting. But there are limits to the productivity of metaphorical expressions in language, and not all neologisms are sensible. A 'high soprano' means one who sings high pitches, and a 'low bass' one who sings low pitches. These expressions are not synonymous with a 'tall soprano' or a 'short bass'; these tall-short expressions would presumably be understood by almost any English speaker as referring to the singers' physical statures, not to their vocal ranges. Likewise, a 'tall sound' does not naturally

mean a high-pitched sound: a listener might interpret this odd expression as meaning a sound that emanates from a high location, or interpret 'tall' as describing some aspect of the sound's timbre. Not only are tall and short never used conventionally to talk about pitch in English, it is difficult to imagine scenarios in which they could even be used sensibly (i.e., comprehensibly) to refer to pitch.

Yet, despite their different lexical profiles and usage restrictions, when high-low and tall-short are used literally, they refer to spatial representations that are similar (see also Dirven & Taylor, 1988; Taylor, 2002) and perhaps overlapping or even identical at the schematic level of spatial representation that is posited to operate in orientational metaphors. Importantly, both pairs of terms refer to points or regions at the top and bottom of a vertical spatial continuum. In some languages, high-low and tall-short are not even distinguished lexically. In Italian, for instance, *alto* can mean high or tall, and *basso* can mean low or short (Goy, 2002), providing one indication that the spatial schemas corresponding to high-low and tall-short may be quite similar.⁵

For the present study we designed a space-pitch congruity task in which English-speaking participants made binary speeded judgments on high-frequency and low-frequency pitches. In one version of the task they classified the pitches according to the familiar linguistic labels (high and low), and in the other they classified them according to novel labels (tall and short). Previous studies have shown reaction time congruity effects in tasks that required participants to classify pitches using the words 'high' and 'low'. Responses were faster when these words were matched to the higher- and lower-frequency pitches, respectively, than when they were mismatched (Melara & Marks, 1990; see also Ben-Artzi & Marks, 1999). We reasoned that if these congruity effects were due to the words high and low activating a vertical spatial schema, then substituting the words tall and short, which cue the same (or a similar) vertical schema should produce similar space-pitch congruity effects, despite the fact that these words cannot be used to describe pitches in English.

⁵ The fact that one language has a single spatial term where another language uses distinct terms does not necessarily mean that the distinct terms refer to exactly the same spatial schema. For example, Spanish uses 'en' to describe containment relations that English speakers would describe as 'in' and support relations that English speakers would describe as 'on' (Bowerman & Choi, 2003).

According to an alternative view of metaphor, however, linguistic metaphors do not cue non-linguistic source-domain representations. Contra Lakoff, some theorists believe the locus of metaphor is in language, and that verbal metaphors are important communicative devices but not clues to conceptual structure (McGlone, 2007; Murphy 1996, 1997). On this view, the reaction times congruity effects observed previously could be due to the polysemy of the words high and low, which refer in some instances to space and in other instances to pitch. RTs are faster when the spatial and musical senses of the words are matched, and slower when they are mismatched. We reasoned that if congruity effects result from the match or mismatch of the senses of polysemous words, then such effects should disappear if the same tasks were performed with the words tall and short, which have no documented auditory or musical senses.

Experiment 1: The wrong words but the right spatial schema

Methods

Participants.

Twenty-five native English speakers with no reported hearing problems participated in exchange for payment (5\$ per 30 minutes). One participant had to be excluded from analyses for not following instructions. In 2 participants technical problems were faced (incomplete outputfiles were produced) and 2 further participants responded according to the wrong response mapping throughout one condition. The latter 4 participants were replaced by a new sample of 4 participants who had not previously participated in the task, resulting in 24 participants.

Materials and Procedure.

Participants were asked to classify tones (one high and one low pitch) as quickly and accurately as possible by pressing buttons on the QWERTY keyboard (R- and O-keys). Stimuli were presented on a pc laptop using presentation software (version 14.9, www.neurobs.com). Sounds were generated by Audacity software (<http://audacity.sourceforge.net/>) and comprised two pure tones

(frequency: 262 and 440 hertz). Each tone lasted 400 ms. Participants listened to one tone at a time, via sealed headphones. Immediately following the offset of each tone, two response options (e.g., high, low) appeared, one on the bottom left and the other on the bottom right of the screen. Participants were instructed to classify the sound by pressing the button located under the corresponding word (e.g., high or low) as fast and accurately as possible. The left-right locations of the spatial terms varied randomly from trial to trial so that participants could not predict the location of the correct word in advance.

Spatial terms (high-low vs. tall-short) were presented in 2 blocks, a high-low block and a tall-short block. Within each block, spatial terms were crossed with 2 mappings (congruent, incongruent). The order of blocks was counterbalanced across participants. The order of congruity was counterbalanced within each block. Across blocks, incongruent and congruent conditions were always presented in alternation. Before each condition, participants received 6 practice trials with feedback. Participants were also given an example illustrating the respective mapping before the practice trials.

Each condition consisted of 24 trials, 96 trials in total. In half of the trials a high pitch was presented, in the other half a low pitch. In the high-low congruent condition, the high pitch had to be classified as *high* and the low pitch as *low*. In the high-low incongruent condition, the high pitch had to be categorized as *low* and the low pitch as *high*. In the tall-short congruent condition the high pitch had to be categorized as *tall* and the low pitch as *short*. In the tall-short incongruent condition the low pitch had to be categorized as *tall* and the high pitch as *short*.

Results

All data were analyzed using R (version 2.14.2; <http://www.r-project.org/>) and the R packages *lme4* (Bates & Maechler, 2009) and *languageR* (Baayen, 2009; cf. Baayen, 2008). We carried out linear mixed-effects regression models of Language (high/low versus tall/short) and Congruency (congruent, incongruent) on accuracy and RTs. Using the principle of backward selection, we started out with a full (conservative) model which took into consideration not only the random intercept but also the random slopes of subject whenever it was appropriate (i.e., when the factor was a within-subject factor). Random

intercepts and slopes of items were not included in the analysis due to the small number of items (4 words: high/low, tall/short). To interpret the significance, we adopted the criterion that a given cosine was significant if the absolute value of the *t*-statistic (or *z*-statistic) exceeded 2 (Baayen, 2008). When appropriate (i.e. when the number of observations exceeded 1000 observations), we also calculated an upper bound for the degrees of freedom by taking the number of observations and subtracting the number of fixed-effects parameters in order to obtain *p*-values (Baayen, 2008).

Accuracy The mean accuracy for all target trials was 95.5% (*SD* = 5.3). For the high/low block, accuracy was 96.1% (*SD* = 5.0), and for the tall/short block, accuracy was 94.9% (*SD* = 6.3). For congruent conditions, accuracy was 95.8% (*SD* = 4.5) and for incongruent conditions it was 95.2% (*SD* = 7.2). Analyzing accuracy by using a logistic mixed effects model on binary accuracy data yielded no main effects or interaction of Language (high/low, tall/short) or Congruency (congruent, incongruent) (Language: $z=|0.1|$; Congruency: $z=|1.3|$; Language by Congruency: $z=|0.1|$).

Reaction times Reaction times for the button presses were analyzed by linear mixed effects models. Only correct trials were considered. This resulted in the exclusion of 5% of the data. Responses greater or less than ± 2 SDs away from each participant's average RTs were also excluded, which resulted in the removal of 5% of the accurate trials.

There were significant main effects of both Congruency ($t=|4.0|$) and Language ($t=|2.0|$), and a significant interaction of Congruency by Language ($t=|2.0|$), $p=0.05$. A linear mixed effect model of Congruency on reaction times restricted to high/low conditions, yielded a significant effect of Congruency ($t=|4.3|$), demonstrating a congruity effect of high/low conditions. Restricting the model to tall/short conditions, resulted in a significant effect of Congruency ($t=|2.3|$), demonstrating a congruity effect of tall/short conditions (see Figure 1).

To further investigate whether congruity effects depended on the order in which participants performed the high/low and tall/short blocks, we included

block order (height first, tallness first) as a between-subjects factor. There was no significant interaction of Language by Congruency by Block order ($t=|0.9|$).

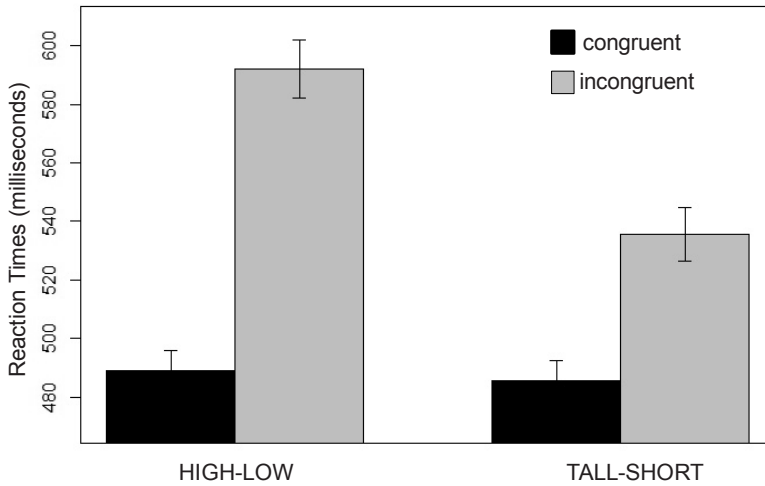


Figure 1: The influence of Language (high-low; tall-short) and Congruency (congruent; incongruent) on pitch categorization (RTs plotted in milliseconds).

We found congruity effects in both high-low and tall-short blocks, suggesting that the wrong words can activate mental metaphors for musical pitch so long as they cue the same (or at least a similar) vertical spatial schema as the right words.

However, high-low and tall-short terms have more in common than their reference to vertical extent. First, abstracting away from their verticality, both pairs of terms name the poles of a linear spatial continuum. Perhaps people will intuit an association between pitch and *any* bipolar linear continuum. Second, both antonym pairs share the same pattern of markedness. Unmarked terms are commonly defined as the default, evaluatively positive or broader term as opposed to the marked one (see e.g. Lehrer, 1985; for a critical approach see Haspelmath, 2006). Both *high* and *tall* would thus be considered unmarked terms and both *low* and *short* marked terms (Clark, 1973). Whereas high can be used to refer to a whole dimension (height), the same is not true for 'lowness'. Similarly, one would describe a building as 'ten meters high', but not as 'ten meters *low' (Clark, Carpenter, Just, 1973). Markedness is not necessarily restricted to linguistic antonyms; it also extends to nonverbal stimuli such as

arrows, which is why some researchers refer to it more generally as polarity (e.g., Proctor & Cho, 2006; Lakens, 2012).

There is evidence that patterns in markedness (or polarity) can affect various binary judgment tasks (e.g., Clark, Carpenter, & Just; 1973; Proctor & Cho, 2006; Lakens, 2012). Responses in a parity judgment task, for instance, were facilitated when stimuli and response codes corresponded in markedness (even–right, odd–left) while incongruent (opposing) conditions (even–left, odd–right) led to interference (Nuerk, Iversen, & Willmes, 2004). In principle, markedness could also underlie the congruity effects we found in Experiment 1. Participants may find it easier when they have to classify the unmarked end of the pitch continuum (high pitch) using an unmarked term (high or tall) and the marked end (low pitch) using a marked term (low or short), compared to the reverse.

Experiment 2 was designed to evaluate these two skeptical interpretations of the results of Experiment 1. We tested a new cohort of native English speakers in the same speeded classification task. However, the words 'short' and 'tall' were replaced with another pair of spatial terms: 'front' (unmarked) and 'back' (marked) (see e.g., Clark, 1973; Landsberg, 1995; Lyons, 1978). If markedness caused the tall-short congruity effect, then we expect to see a similar congruity effect for front-back. Likewise, front-back should produce a space-pitch congruity if these effects can be caused by *any* pair of spatial terms (or at least any terms that describe a bipolar linear continuum, regardless of its orientation). If however, the congruity effects in Experiment 1 were due to words activating a vertical spatial schema, then we expect to see no congruity effect in the front-back condition.

Experiment 2: Will any marked spatial continuum suffice?

Methods

Participants.

Twenty-four native English speakers with no reported hearing problems participated for payment (5\$ per 30 minutes). Four participants were excluded from analyses for not following instructions (i.e. they responded according to the wrong response mapping throughout at least one condition). They were replaced

by a new sample of 4 participants who had not previously participated in the task.

Materials and Procedure The same procedure as in Experiment 1 was used, with the following exceptions. Rather than classifying pitches as high-low or short-tall, participants classified them as high-low for one block and *front-back* for the other. Stimuli were presented on an Apple iMac using Vision Egg 2.6 (Straw, 2008). Participants were asked to press buttons on the keyboard (Q- and P-button) in order to classify the tones.

In the tall-short congruent condition the high pitch had to be categorized as *tall* and the low pitch as *short*. In the tall-short incongruent condition the low pitch had to be categorized as *tall* and the high pitch as *short*. In the front-back congruent condition the high pitch had to be categorized as *front* and the low pitch as *back* (according to patterns of markedness). In the front-back incongruent condition the low pitch had to be categorized as *front* and the high pitch as *back*.

Results

All data were analyzed using R (version 2.14.2; <http://www.r-project.org/>) and the R packages *lme4* (Bates & Maechler, 2009) and *languageR* (Baayen, 2009; cf. Baayen, 2008). We carried out linear mixed-effects regression models of Language (high-low versus front-back) and Congruency (congruent, incongruent) on accuracy and RTs. Using the principle of backward selection, we again started out with a full (conservative) model which took into consideration not only the random intercept but also the random slopes of subject whenever it was appropriate (i.e., when the factor was a within-subject factor). Random intercepts and slopes of items were not included in the analysis due to the small number of items (4 words: high/low, front/back). To interpret the significance, we adopted the criterion that a given cosine was significant if the absolute value of the *t*-statistic (or *z*-statistic) exceeded 2 (Baayen, 2008). When appropriate (i.e. when the number of observations exceeded 1000 observations), we also calculated an upper bound for the degrees of freedom by taking the number of

observations and subtracting the number of fixed-effects parameters in order to obtain p-values (Baayen, 2008).

Accuracy The mean accuracy for all target trials was 92.4% ($SD = 8.1$). For high/low conditions, accuracy was 92.4% ($SD = 9.7$), and for front/back conditions, accuracy was 92.5% ($SD = 11.3$). For congruent conditions, accuracy was 95.9% ($SD = 4.5$) and for incongruent conditions it was 88.9% ($SD = 15.5$). Analyzing accuracy by using a logistic mixed effects model on binary accuracy data yielded no main effects or interaction of Language (high/low, front/back) and Congruency (congruent, incongruent), (Language: $z = 1.3$; Congruency: $z = 0.2$; Language by Congruency: $z = 1.2$).

Reaction times Reaction times of the button presses were analyzed by linear mixed effects models. Only correct trials were considered which resulted in the exclusion of 7% of the data. Responses greater or less than ± 2 SDs away from each participant's average RTs were also excluded, which resulted in the removal of 6% of the accurate trials.

There was no significant main effects of Language ($t = 0.3$). The model yielded a significant main effect of Congruency ($t = 3.5$) and a significant interaction of Congruency by Language ($t = 3.3$). A linear mixed effect model of Congruency on reaction times restricted to the level of high/low, yielded a significant effect of Congruency ($t = 4.5$), demonstrating a congruity effect of high/low conditions. Restricting the model to the level of front-back yielded no significant effect of Congruency ($t = 0.2$) (see Figure 2).

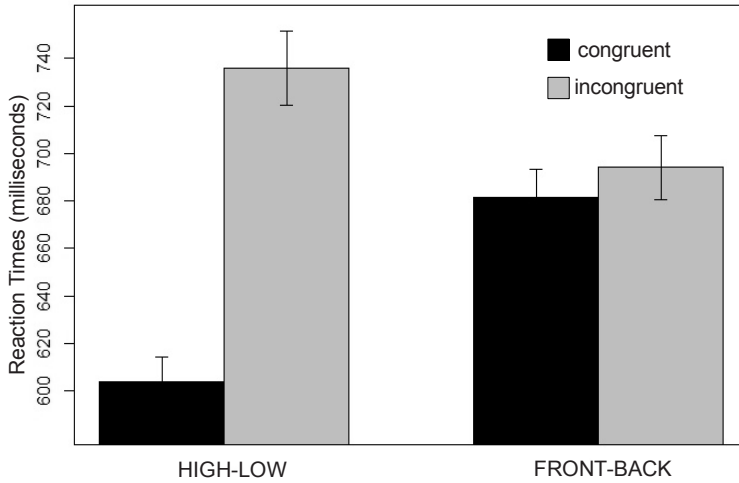


Figure 2: The influence of Language (high-low; front-back) and Congruency (congruent; incongruent) on pitch categorization (plotted in milliseconds).

Between experiment comparison

In both Experiment 1 and 2, English speaking participants responded faster when pitch stimuli had to be classified as being high or low (in line with their familiar height-pitch metaphors). Moreover, participants also showed a congruity effect for tall-short conditions, but not for front-back condition. In order to test whether congruity effects differed between the two experiments, we compared the two in a mixed effects model. Overall there was a 3-way interaction of Language by Congruency by Experiment ($t=|2.6|$). A further analysis compared the magnitude of the congruity effects for tall-short versus front-back. A significant interaction of Congruency by Experiment ($t=|2.5|$) indicated that the congruity effect was greater for tall-short than for front-back.

General Discussion

We have shown that alongside congruity effects in high-low conditions, people also respond faster to congruent tall-short conditions compared to incongruent conditions in pitch classification tasks. This finding suggests that wrong words can activate height-pitch metaphors so long as they cue the

relevant spatial schema of verticality. Words that activate a different (irrelevant) spatial schema (front-back), however, do not result in congruity effects. This finding further indicates that congruity effects cannot be attributed to markedness (polarity alignment), since 'front' and 'back' also name the unmarked and marked ends of a (sagittally oriented) spatial continuum.

However, given that high-low and tall-short are closely related, it is possible that congruity effects for the words 'tall' and 'short' arose either due to strategic activation of the words 'high' and 'low', or to semantic priming from these words. Although these possibilities cannot be ruled out entirely, a number of facts speak against them. First, presumably participants would be more likely to activate 'high' and 'low' in the tall-short condition when this condition was preceded by the high-low condition, yet the strength of the tall-short congruity effect did not depend on the order of blocks (i.e., there was no interaction between Congruency by Block order). Simple spreading activation from 'high' to 'tall' and 'low' to 'short' is also unlikely to account for the tall-short congruity effects. According to Latent Semantic Analysis (LSA; <http://lsa.colorado.edu/>), 'tall' is more strongly related to 'short' (LSA cosine: .48) than to 'high' (LSA cosine: .31). Moreover, 'short' is about equally strongly related to 'high' (LSA cosine: .30) as to 'low' (LSA cosine: .31). Since activation is expected to spread between the most strongly related items (Collins & Loftus, 1975), simple spreading activation would have wiped out a tall-short congruity effect rather than producing it.

Both tall-short and high-low words were found to activate mental metaphors for pitch. However, the congruity effect in tall-short condition was weaker than in high-low condition. This could be because high-low and tall-short activate slightly different spatial schemas. While both pairs of spatial terms activate a schematic representation of verticality, tall-short seems to be associated with 3-dimensional size information, as well. For instance, a tall person is also likely to be larger than a short person, all around (e.g., Clark, Carpenter, & Just, 1973). Unlike the alignment of spatial height and pitch (high = high pitch), however, links between size and pitch have been shown to follow a reverse direction: 'big' is associated with low pitch, and 'small' with high pitch (e.g., Bien, ten Oever, Goebel, & Sack, 2012; Evans & Treisman, 2010; Parise &

Spence, 2009). Competing schemas associated with tall-short terminology (i.e., 1-dimensional verticality versus 3-dimensional size) may thus have led to weaker congruity effects in tall-short as compared to high-low blocks.

Overall, however, we found that height-pitch mappings can be activated by 'tall' and 'short'. This finding is at odds with theories proposing that the locus of metaphor is in language (e.g., McGlone, 2007). Congruity effects are not restricted to polysemous high-low terminology with overlapping terms for height and pitch but extend to spatial terms with no musical senses. It thus seems that the 'active ingredient' of space-pitch metaphors lies outside of language (i.e. in the activation of non-linguistic spatial schemas). In accordance with this proposal, height-pitch mappings can be found in prelinguistic infants and in nonlinguistic tasks, as well as when incorrect language is used. Whereas language can play a role in activating space-pitch mappings, thereby strengthening them, it seems that the locus of space-pitch metaphors is in thought rather than in language.

Acknowledgments

We would like to thank Tom Gijssels for programming the Python version of the experiment and for help with the experimental set-up. Also, we would like to thank Cleve Graver for his assistance in data collection and Geoff Brookshire for help with R-analyses.

Searching high and low: The relation of space and musical pitch in the brain

Chapter 5

Based on:

Dolscheid, S., Willems, R.M., Hagoort, P., & Casasanto, D. (in preparation).
Searching high and low: The relation of space and musical pitch in the brain

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 completed spatial judgment tasks in two modalities; vision and touch. They also compared tones that varied in pitch. 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 and tactile areas, suggesting that space-pitch associations may involve modality-specific spatial representations.

Introduction

According to metaphor theories, many of our abstract concepts are spatial in nature (Lakoff & Johnson, 1980). Indeed, it was shown that spatial representations contribute to people's understanding of numerous domains, including time (Boroditsky, 2000), preference (Casasanto, 2009), and social dominance (Schubert, 2005). Unlike these abstract cases, however, spatial representations also seem to underlie domains that can be perceived directly, like musical pitch.

In a seminal study by Pratt (1930) participants were asked to locate the position of different pitches on a numbered scale running from floor to ceiling. Tones were generated by a hidden loudspeaker varying in vertical position. Independently from the actual location at which the tones were played, participants consistently assigned higher pitches to higher positions and conversely lower pitches to lower positions in space (Pratt, 1930). In more recent speeded classification tasks, people responded faster to a stimulus when visuospatial height and pitch were congruent as compared to incongruent stimulus presentation (Melara & O'Brien, 1987, see also Ben-Artzi & Marks, 1995; 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 was 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. Three- to 4-month-olds look significantly longer in a preferential looking task at congruent trials (in which visuospatial height and pitch height are aligned) as compared to incongruent trials (Walker et al., 2010, see also Dolscheid, Hunnius, Casasanto, & Majid, 2012; Wagner, Winner, Cicchetti, & Gardner, 1981).

Although numerous behavioral experiments confirm links between space and musical pitch, they do not shed light on the neuronal underpinnings of space-pitch associations. To what extent do pitch representations rely on neural circuitry that is involved in spatial processing? And what is the exact nature of the spatial representations underlying musical pitch?

Theories of embodied cognition suggest that representations (like space or motion) are grounded in systems of perception and action, and are captured by conjunctive neurons in cortical association areas. In a process called re-enactment or simulation, these conjunctive neurons activate sensory-specific neuronal populations and thereby partially re-enact perceptual states in the absence of sensory input (Barsalou, Simmons, Barbey & Wilson, 2003). Whereas modality-specific re-enactment has been demonstrated for a number of domains (e.g., motion, Saygin, McCullough, Alac, & Emmorey, 2010), little is known about simulation in metaphorical processing. Are spatial representations that underlie other domains like musical pitch modality-specific in nature? More precisely, does musical pitch rely, at least in part, on visuospatial representations, as has been suggested by some researchers (see for instance 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; Merrill et al., 2012; Perry et al., 1999; Zatorre et al., 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 visuospatial in nature.

Crucially, however, space is not restricted to visual perception. Rather, spatial experiences are inherently 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 (Renier et al., 2009; Brandt & Dietrich, 1999). These

regions are involved, for instance, when participants have to determine the type of spatial relation between object pairs (e.g., above/below) (Amorapanth, Widick, & Chatterjee, 2010; Corradi-Dell-Aqua, Hesse, Rumati, & Fink, 2008; Quadflieg et al., 2011).

However, activations in regions like IPS are not restricted to spatial judgments but also extend to pitch-related tasks. Schwenzer and Mathiak (2011), for instance, reported IPS activations 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. 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.

However, since no study to date has directly compared the neural correlates of spatial height and pitch height, the nature of space-pitch associations cannot be determined. Are spatial representations of pitch unimodal (e.g., visuospatial) in nature or are they multimodal, or maybe both? In order to answer these questions, we investigated the spatial basis of pitch representations in an fMRI experiment. To determine whether pitch representations overlap with unimodal or multimodal spatial representations, we asked participants to judge spatial height in two modalities; vision and touch. In order to localize regions involved in spatial height processing, 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, whereas in the tactile condition, participants felt the stimuli on their palms. In addition, participants were asked to judge whether two sounds differed in pitch. We reasoned that if pitch judgments rely on the same neural circuitry as judgments concerning visual, tactile, or multimodal space (the intersection of visual and tactile space), we should see pitch activations in region(s) that are involved in spatial height processing.

In order 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. Participants were presented with the same stimuli as in the critical conditions, but rather than judging whether the stimuli were the same height, in the control conditions participants judged whether two successive stimuli were the same shape. Focusing attention on one dimension (e.g., space) should enhance cerebral activity associated with the processing of this dimension, a principle that has been proven to be valid in a number of sensory domains (e.g., Roland & Friberg, 1985; Corbetta et al., 1990; Platel et al. 1997). Subtracting activations of the control condition (shape) from spatial activations (height) should thus result in activations that are primarily indicative of visual and tactile spatial height processing. 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 – as a control – the same timbre (i.e., played by same or different instruments). Subtracting activations of the control condition (timbre judgments) from pitch judgment activations should result in activations that are primarily indicative of musical pitch (see also Platel et al., 1997).

If pitch representations are instantiated in modality-specific visual cortices, pitch judgments should activate areas that are also activated by visuospatial height processing. Alternatively, if pitch representations draw on modality-specific cortices that are not restricted to vision, pitch judgments should activate areas activated by tactile height processing. If pitch representations are based on multimodal space, however, pitch judgments should activate regions that are activated by both visual and tactile spatial (multimodal) height judgments. Finally, it is possible that pitch judgments could activate both modality-specific and multimodal spatial areas (i.e., both the common and distinctive regions activated by visual and tactile height judgments).

On the basis of these predictions, we also defined several regions of interest (ROIs). We looked for potential pitch activations in unimodal and multimodal ROIs. Primary visual cortex (BA 17) served as a visual ROI (Noesselt et al., 2002). The postcentral gyrus, a region that is 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 was lost due to storage problems. Three participants had to be excluded from further analyses due to large head movement during the scanning session. The study was approved by the local ethical committee for research with human participants.

Materials.

Three different types of materials were used in three corresponding 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 participants' palms, either at a position high in tactile space (upper part close to the fingers of the participants) 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).

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 comparably low-pitched sound (262 hz) and a comparably high-pitched sound (394 hz).

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 (Figure 1). 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 and response times were recorded. Overall, two dimensions of shape (circle and square) were fully crossed with two dimensions of position (high and low). This means that either shape was presented in the upper or in the lower position on the screen. The resulting 4 different stimulus options could take up both the first and the second position in the stimulus sequence, resulting in 16 different combinations. In total, the visual

block consisted of 64 trials, 32 position trials and 32 shape (control) trials. The order of stimuli 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 (Picture 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 finger, low=close to the participant's wrist). In total, the tactile block consisted of 64 trials, 32 position trials and 32 shape (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 pitch trials and 32 timbre (control) trials.

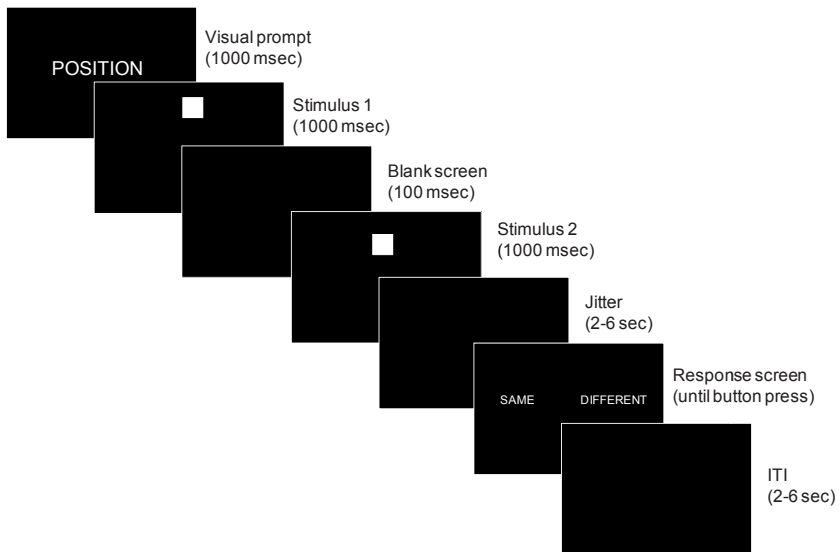
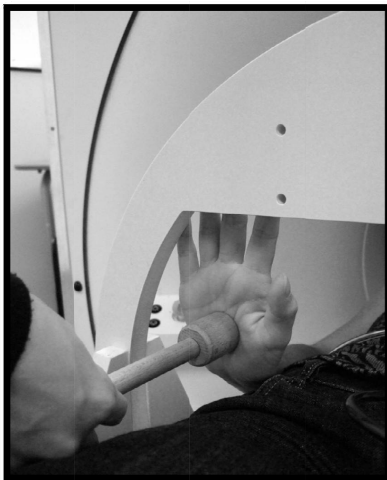


Figure 1: Examples of a visual trial. Participants saw a visual prompt at the beginning of each trial. Then two stimuli were presented. Participants were asked to compare these stimuli with respect to the relevant dimension (in this example; position). After a jitter, participants responded with their left hand indicating whether the presented stimuli differed or not. The same trial structure was also employed in the tactile as well as the auditory block.



Picture 1: Illustration of the 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).

Behavioral data

Number of errors and response times (RTs) for the judgment tasks were collected.

fMRI Data acquisition

Functional data were acquired on a Siemens Avanto 1.5 T MRI system (Siemens, Erlangen, Germany) using a standard birdcage head-coil for RF transmission and signal reception. T2*- weighted BOLD-sensitive images were acquired using a gradient EPI sequence (Echo Time = 35ms; Repetition Time = 2.34s; 32 axial slices in ascending order with a 10% gap between slices; voxel size = 3.3 x 3.3 x 3.5 mm). For each subject we also acquired a T1-weighted high-resolution anatomical scan (Echo Time = 2.95 ms, Repetition Time = 2.25s, voxel size = 1.0 x 1.0 x 1.0 mm, 176 sagittal slices).

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 were entered into a General Linear Model with separate regressors for each of the conditions (visual height, visual shape; tactile height, tactile shape; auditory pitch, auditory timbre). Regressors were modeled for the length of a trial (duration of subsequent stimuli) and were convolved with a canonical hemodynamic response function

(HRF). Only trials with a correct response were considered. 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 movements. Since we were interested in activations that could be exclusively attributed to spatial processing, we subtracted activations of shape (control) from activations of spatial height ([height-shape]) for both the visual and the tactile modality. In order to reveal exclusively pitch related activation, we applied the same method to auditory stimuli ([pitch-timbre]).

A second-level whole-brain group analysis with subjects as a random factor ("random effects analysis", Friston et al., 1999) was carried out. Here, we looked for regions that were selectively activated by spatial height ([height-shape]) in vision and touch, as well as multimodal regions activated by both visual and tactile height [visual (height-shape) \cap tactile (height-shape)]. Conjunction analyses were conducted by using the Minimum Statistic compared to the Conjunction Null (Nichols, Brett, Andersson, Wager, & Poline, 2005). We also looked for regions that were selectively activated by pitch ([pitch-timbre]). 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 regions of interest analyses. In the first ROI analysis, we looked for pitch activations ([pitch-timbre]) and ([pitch]) in regions that were selectively activated by visual height, tactile height and multimodal height. Voxels with a $p < .001$ (whole brain, uncorrected) were taken as ROIs.

In a second ROI analysis we looked at visual as well as tactile ([height-shape]) activations in predefined anatomical regions of interest. In case of significant visual and/or tactile ([height-shape]) activations, we also looked for pitch activations ([pitch-timbre]) and ([pitch]) 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 given 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 as well as its subregions (left angular gyrus, left supramarginal gyrus) were selected as multimodal ROIs (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%, SD = .08). In the visual task, performance was high for both the height (Mean accuracy = 97%, SD = .04) and the shape (Mean accuracy = 97%, SD = .07) conditions. In the auditory task, performance was high for both the pitch (Mean accuracy = 92%, SD = .07) and the timbre (Mean accuracy = 94%, SD = .07) judgments. In the tactile task, performance was high for the height (Mean accuracy = 97%, SD = .03) but slightly lower in the shape (Mean accuracy = 78%, SD = .12) condition.

Comparing the numbers of errors during the different modality blocks, a repeated-measures ANOVA revealed a significant effect of modality, $F(2,30)=15.99$, $p=.0001$, $\eta_p^2=.52$. Pairwise comparisons showed that participants produced significantly more errors in the tactile block as compared to the auditory block ($p=.004$, Bonferroni corrected). Participants also produced significantly more errors in the tactile block as compared to the visual condition ($p=.001$, Bonferroni corrected). The number of errors between the visual and the auditory condition did not differ statistically.

Since participants mainly reported difficulties with the tactile shape recognition task as also indicated by the lower accuracy rate, we compared errors in the tactile shape condition with errors in the tactile space condition. A paired-sample t-test revealed that participants produced significantly more

errors in the tactile shape as compared to the space condition, $t(15)=7.03$, $p=.0001$. Since tactile shape judgments served as a control condition and only trials with a correct response were included in the fMRI analysis, higher error rates were not considered problematic.

A repeated-measures ANOVA revealed no significant differences of response times between the three different modalities (vision, audition, touch), $F(2,30)=1.33$, ns , $\eta_p^2=.08$.

Whole Brain analyses

Spatial activations: A conjunction of visual ([height-shape]) and tactile ([height-shape]) resulted in no activations that survived the statistical threshold. Informal inspection revealed that likewise there were no activations at $p<0.001$ uncorrected. There were no significant activations of visual ([height-shape]) nor of tactile ([height-shape]).

Pitch activations: Significant activations of ([pitch-timbre]) were found in the thalamus [MNI coordinates: 6, -6, 18].

Anatomical regions of interest (ROI) analyses

Visual: There was no significant activation of visual ([height-shape]) in the visual ROI (BA 17).

Tactile: The ROI analysis revealed significant activations of tactile ([height-shape]) in the left postcentral gyrus [MNI coordinates: -66, -20, 26] (Figure 2). We used this region as a region of interest (ROI) and found significant activity of ([pitch]) ($t(15)=4.23$, $p=.0001$). However there were no significant activations of ([pitch-timbre]), $t(15)=.77$, ns .

Importantly, the reported postcentral cluster was not affected by an exclusion mask ($p=.01$) of visual height, indicating that activity in this region can mainly be attributed to (unimodal) tactile height and not to visual height.

Multimodal: There was no significant activation of the conjunction [visual (height-shape) \cap tactile (height-shape)] in the IPL.

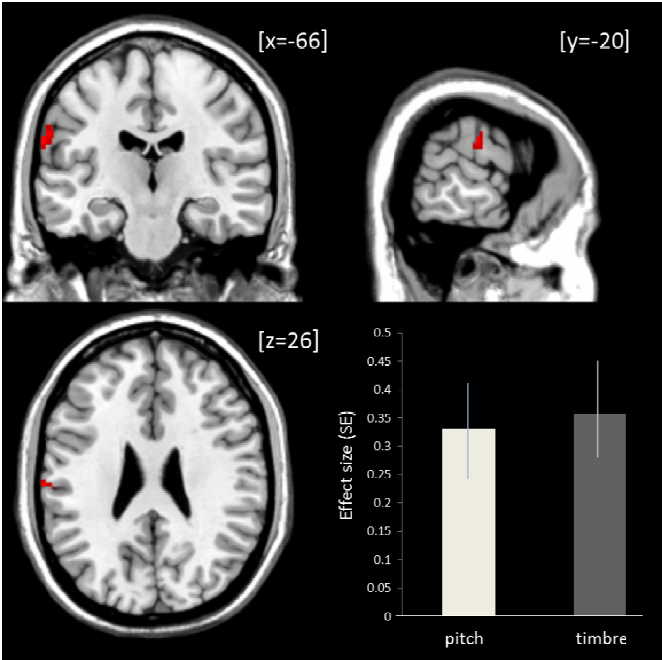


Figure 2: Significant tactile [(height-shape)] activations in postcentral gyrus. In addition to tactile height sensitivity, there was also significant pitch activity, but no effect of [(pitch-timbre)].

Discussion

In whole brain analyses, we did not find any brain areas that were significantly more active during spatial height judgments than during shape judgments. This was true for both the visual and the tactile modality. It appears that our manipulation was too subtle to elicit differential patterns of activation. This finding is surprising, given that the behavioral data indicate that participants were attending to height during 'height' trials and to shape during 'shape' trials. This finding is consistent, however, with evidence indicating that the classical distinction between ventral ('What' e.g. shape) and dorsal ('Where' e.g. space) streams of processing (Haxby, Grady, Horwitz, Ungerleider, &

Mishkin, 1991; Ungerleider & Haxby, 1994) seems to be more dynamic and less strict than previously thought (e.g., Koshino et al., 2005).

Likewise, except for activity in the thalamus, a subcortical region involved in pitch processing (e.g., Wong, Parsons, Martinez, & Diehl, 2004), there were no cortical activations selectively responding to pitch as compared to timbre. This lack of selective activation is at odds with previous findings (e.g., Platel et al., 1997). A potential explanation could be that the perceptual attributes of pitch and timbre are so-called integral dimensions (Garner, 1974; Melara & Marks, 1990). Unlike separable dimensions in which attention can be selectively allocated to the relevant stimulus attribute (e.g., phoneme versus pitch; Day & Wood, 1972), integral dimensions are intertwined such that categorizing a stimulus according to either dimension (e.g. timbre) is also influenced by variation on the irrelevant dimension (e.g. pitch). Since behavioral results suggest that it is hard to ignore timbre information while making judgments about pitch, this might explain why we did not find any differential effect of attention to musical pitch versus timbre at a cortical level. That is, although participants were clearly attending to the relevant dimension of the stimuli, they may also have been attending inadvertently to the irrelevant dimension. Pitch information is processed in a highly automatic fashion, perhaps even when participants are instructed to attend to timbre; as such, 'selective attention' to one dimension of tone or the other may not be selective enough to yield significant differences when pitch and timbre conditions are compared directly (see also Zatorre, 1999).

We also did not find any significant activity of multimodal space [visual (height-shape) \cap tactile (height-shape)], neither in the whole brain nor in predefined ROIs. However, we did find significant tactile activations [(height-shape)] in the left postcentral gyrus. Additionally, the analysis revealed significant activations of pitch in this activated ROI, suggesting that there might be a unimodal tactile basis for musical pitch.

Unlike a number of findings that reported pitch-related activity in primary and secondary visual areas as well as in multimodal areas, here we did not find any evidence for pitch activations in these regions. Maybe the absence of these effects was caused by specifics of our experimental design. Since shapes as

well as timbres were fully crossed with spatial positions and pitch 'positions', this means that in half of the trials spatial and pitch height actually changed (from low to high or vice versa) whereas in the other half the position of the stimulus remained the same. Collapsing over instances of constant and varying position might thus have washed out effects. Looking instead separately at trials in which position (of space and pitch) changed versus trials in which they remained constant might give us a better insight into overlap between the dimensions of interest. We therefore modeled the trials differently, comparing activity during trials for which space or pitch height changed to activity during trials where height or pitch stayed the same (control trials). Since changes in position could only be determined at the time the second stimulus was presented, we only included the second stimulus in our model.

FMRI data analyses II

The time series of the preprocessed data were entered into a General Linear Model with separate regressors for the stimulus that was presented secondly. This time, 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), resulting 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. Regressors 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) \cap tactile (height change-control)]. We also looked for regions that were selectively activated by pitch change ([pitch change-control]). To correct for multiple comparisons, 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 actual height perception, we performed two regions of interest analyses. In the first ROI analysis we looked for activations of ([pitch change-control]) and ([pitch change]) 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-control]) and ([pitch change]) in the activated region(s). ROIs again comprised BA 17, the postcentral gyrus as well as the IPL.

Results of height change analyses

Whole brain analyses and functional ROI analyses II

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 of ([pitch change-control]), $t(15)=1.41, ns$.

Moreover, this cluster revealed a significant deactivation in tactile height change ([height change]) ($t(15)=-5.22, p=.0001$) but no significant activations of tactile ([height change-control]), $t(15)=-.28, ns$.

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) \cap tactile (height change-control)].

Auditory activations: There were no significant activations of [(pitch change-control)].

Anatomical ROI analyses of height change

Visual: A ROI analysis in BA 17 (anatomically defined) revealed a significant cluster of visual [(height change-control)] activity [MNI coordinates: 2, -84, 10] (Figure 3). Within this region pitch change led to significantly more activation as compared to the auditory control condition [(pitch change-control)], $t(15)=2.17$, $p=.05$. There was significant activation during pitch change ($t(15)=3.4$, $p=.004$), whereas the auditory control condition revealed no significant activity ($t(15)=1.6$, ns). Moreover, this cluster revealed a significant deactivation in tactile height change ($t(15)=-5.26$, $p=.0001$) but no significant activations of 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) \cap tactile (height change-control)] in the IPL.

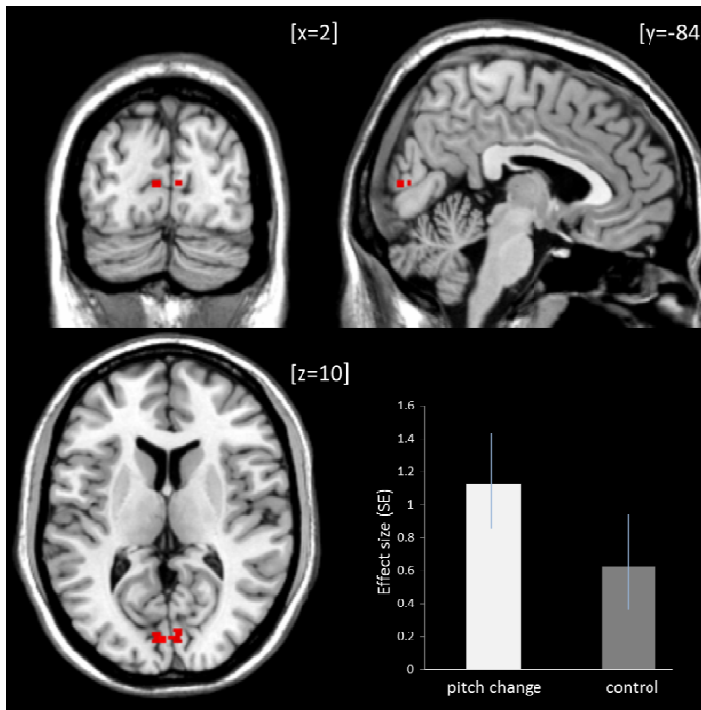


Figure 3: Significant visual [(height change-control)] activity in BA 17. In addition to visual height sensitivity, there was also significant activity of pitch change. This ROI also revealed significant activations of [(pitch change-control)].

Discussion of height change analyses

By comparing stimuli that differed in spatial height to those that remained at a constant position (control), we found activations in primary visual cortex (BA 17). In this primary visual area we also observed activity of [(pitch change)] however no significant activation of [(pitch change-control)]. Finally, in the ROI analysis, a cluster restricted to BA17 specifically responded to visual [(height change-control)] also revealed significant activity of [(pitch change)] as well as [(pitch change-control)], suggesting a specific overlap between pitch as well as 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 ROI, there was no effect of tactile [(height change-control)], indicating that this

ROI more specifically responded to changes of (visuo-)spatial height and auditory pitch.

General Discussion

Does processing pitch height activate parts of the cortex that are involved in processing spatial height? Here we arrive at a preliminary answer to this question. 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). In line with previous research this activity may reflect cortical sensitivity to visuospatial variation (e.g., Bosking, Crowley, & Fitzpatrick, 2002). We used this area of BA 17 as a visuospatial ROI in which to search for pitch-related activity. Crucially, we observed activity of 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 visual regions during pitch processing has been observed previously. In fact, various other studies have also reported activity of primary and secondary visual areas during pitch-related tasks (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. Pitch judgments seem to rely, in part, on the same neural tissues that support visual perception, as suggested by theories of metaphor and embodied cognition.

It is possible that more general auditory information – not restricted to musical pitch – causes activity in the visual cortex. Romei and colleagues for instance found that looming sounds can selectively enhance the excitability of visual cortex (Romei, Murray, Cappe, & Thut, 2009). Zangenehpur and Zatorre (2010) further demonstrated that BA 17 was recruited by auditory stimulation after repeated coupling of an auditory stimulus with spatiotemporally matched visual stimuli. These findings indicate that vision and sound (beyond pitch) are tightly linked, suggesting stronger connections between visual and auditory areas than previously claimed (see e.g., Romei et al., 2009; Ghazanfar et al., 2005).

Yet, our results show activity of pitch in regions of the visual cortex that are selectively sensitive to variation of spatial positions. This finding suggests a more specific connection between visuospatial and auditory processing, not just some general link between vision and audition. In support of this argument, we find a cluster in BA 17 that is selectively activated by pitch change as compared to control stimuli (see also Figure 3). This area of visual cortex that is selective for processing changes in (visuo-)spatial height (as opposed to any visual stimulus) is also selective for processing changes in pitch height (as opposed to any auditory stimulus).

In addition to pitch-related activity in primary visual areas, we also found pitch-related activity in primary tactile areas, which were active during spatial height judgments (postcentral gyrus). This overlap of activations suggests that modality-specific activations of pitch are not necessarily restricted to the visual domain. However, this finding has to be interpreted with caution. Since we did not find stronger activations of pitch as compared to control conditions, postcentral activations may also indicate more general processes, such as auditory attention, rather than selective processing of pitch height (see Zatorre, Mondor, & Evans, 1996).

These data provide initial evidence for unimodal cortical areas supporting the metaphorical link between space and musical pitch. Yet, much remains to be learned about this cross-domain linkage. One question that remains is why we did not see any selective activations of pitch in multisensory parietal regions. Both the IPS and the IPL play an important role in general attentional processing and participate in the maintenance and control of selective attention (Näätänen & Alho, 2004, Schultz & Lennert, 2009, Wojciulik & Kanwisher, 1999). Since modulation of attention was required in all of our experimental conditions and fairly equal across the different tasks, this might explain why we did not see any differential activation in multisensory regions.

In other tasks, however, the IPL was found to be involved in pitch memory (Rinne, Koistinen, Salonen & Alho, 2009) as well as pitch production (Peck et al., 2009). IPS activity was reported in pitch identification (Schwenzer & Mathiak, 2011) as well as in pitch transformation tasks (Zatorre et al., 2010). While these studies might indicate a link between spatial and pitch processing in

multisensory regions, they also point to potential differences among space-pitch relations. Presumably, mechanisms that underlie pitch processing differ depending on the complexity of the task. Simple pitch comparisons differ from melody transformation (Zatorre et al., 2010) or complex pitch memory (Rinne, Koistinen, Salonen & Alho, 2009). In our task participants were simply asked to compare different pitches, without a complex pitch-related task.

Moreover, pitch representations have been found to be similar to other types of magnitude representations (e.g. number) and likewise evoke the so-called distance effect: The larger the difference between the compared stimuli, i.e. between two different pitches, the shorter the reaction time in pitch comparison tasks (Cohen Kadosh et al., 2008). Here, we presented two different tones that were only one fifth apart. A larger interval or a parametric modulation of pitch distance thus might have resulted in stronger pitch-related activity in the IPL.

So far, evidence for modality-specific activity corresponding to metaphorical source domains has been elusive (see Willems & Casasanto, 2011). For instance, some rather controversial outcomes have been reported in the domain of action-related idiomatic language. While Boulenger, Hauk, and Pulvermüller (2008) found somatotopic activation in motor cortex when participants read idiomatic sentences like 'He kicked the habit' (meaning 'to quit'), other studies have shown that motor regions were only activated when participants heard isolated action verbs ('kick') or literal action sentences ('kick the ball'), but not idiomatic sentences ('kick the bucket' meaning 'to die'; Raposo Moss, Stamatakis, & Tyler, 2009; see also Aziz-Zadeh, et al., 2006). Moreover, a study by Quadflieg et al. (2011) demonstrated that spatial height representations seem to underlie metaphors of valence. However, since patterns of co-activation were restricted to multimodal areas (IPL), Quadflieg et al.'s results are not informative about modality-specific activity of space.

By contrast, our results shed light on modality-specific neural bases of space-pitch associations. Judgments of musical pitch seem to depend in part on visual areas (and perhaps on tactile areas) that are involved in spatial height processing. Modality-specific representations of spatial height thus appear to

contribute to musical pitch processing, confirming a core assumption of embodied theories of metaphor.

Acknowledgments

We thank Paul Gaalman for assistance during the scanning sessions and help with the experimental set-up. Moreover, we thank Kirsten Weber and Florian Krause for assistance during some scanning sessions. We thank Henning Verlage for help with the auditory stimuli and Peter Dolscheid for constructing the tactile stimuli.

General Discussion and Conclusions

Chapter 6

In this dissertation, I investigated relationships between nonlinguistic space-pitch associations and space-pitch metaphors in language. Whereas effects of language on thought have been demonstrated in a number of domains, here I provide the first evidence that language can also have an impact on nonlinguistic representations of musical pitch. Repeated use of linguistic metaphors appears to influence pitch representations (chapter 2), thus supporting the hypothesis that language plays a role in forming space-pitch associations (e.g., Martino & Marks, 1999; Melara & Marks, 1990). Effects of spatial height on pitch persisted under verbal interference (chapter 2), by contrast with other effects of linguistic relativity, indicating that the observed effects of language on pitch were not due to the use of language online during the experimental task. It seems that space-pitch associations are not inherently linguistic; rather, words activate associations between non-linguistic representations of space and pitch (in line with suggestions by Lakoff, 1993). In support of this proposal, I show that even lexically inappropriate spatial words can activate space-pitch associations, so long as the 'wrong' words cue the right spatial schema (chapter 4). Furthermore, links between space and pitch appear to be instantiated, in part, outside of the brain's core language system, in visual areas that process spatial height (chapter 5).

Space-pitch associations are present prior to the acquisition of language (chapter 3), suggesting that language does not establish links between space and pitch in the first place. Still, repeatedly activating a particular space-pitch association while producing and comprehending a particular linguistic metaphor can strengthen this association at the expense of competitors, giving rise to habits of metaphorical mental representation that differ between members of different language groups.

This process by which the use of linguistic metaphors shapes prelinguistically-established nonlinguistic associations was first proposed by Casasanto (2008) to account for cross-linguistic differences in mental metaphors linking space and time. Together, the studies in this thesis provide the most comprehensive evidence for this process to date.

Here I summarize and discuss the main findings of this dissertation. Moreover, I discuss further questions that came up along the way.

Summary

Language can influence nonlinguistic pitch representations

In order to find out whether metaphors in language can have an impact on thought, I took advantage of cross-linguistic diversity in pitch vocabulary. Whereas Dutch speakers talk about pitch in terms of spatial height, Farsi speakers encode pitch in terms of thickness. Differences in language were also reflected in differences in performance on two nonlinguistic psychophysical pitch reproduction tasks. Dutch speakers' pitch estimates (i.e. singing responses) were significantly modulated by spatial height but not by thickness. Conversely, Farsi speakers' pitch estimates were modulated by spatial thickness but not by height. Dutch and Farsi speakers thus seem to represent pitch differently, according to the languages they speak. Moreover, after Dutch speakers were trained to use Farsi-like thickness metaphors, their performance in the nonlinguistic pitch reproduction task resembled that of Farsi speakers, indicating that language can play a causal role in shaping space-pitch mappings. Chapter 2 thus provides evidence that metaphors in language can have an impact on people's nonlinguistic space-pitch representations (for comparable results in space-time metaphors, see Casasanto, 2010).

But *how* does language affect space-pitch associations? To further explore the nature of linguistic influences, Dutch speaker's height-pitch associations were tested under verbal interference (chapter 2). Unlike other linguistic relativity effects (e.g., in the color domain; Winawer et al., 2007), effects of spatial height on pitch persisted when participants were busy with the rehearsal of letter strings (see Connell, Cai, & Holler, 2012, for converging evidence). I thus

propose that the results of these experiments cannot be attributed to an online use of verbal labels but rather emerged from analog relationships between space and pitch in long-term memory, which are partly conditioned by language.

Prelinguistic origins of space-pitch mappings

Rather than simply echoing space-pitch mappings, I found that language can have an impact on nonlinguistic associations. People who use different space-pitch metaphors in language also think about pitch differently. While linguistic metaphors thus seem to be more than mere 'surface realisations' of underlying mappings (Lakoff, 1993), my results do not rule out that space-pitch representations may exist outside the realm of language. In support of this latter proposal, previous findings suggest that height-pitch associations may be established prior to language acquisition (Walker et al., 2010; Wagner et al., 1981). However, not much is known about the origins of other space-pitch associations, such as the relationship between thickness and pitch. In chapter 3 I investigated both thickness-pitch as well as height-pitch associations in prelinguistic infants. My results reveal that 4-month-old babies are sensitive to correspondences between auditory pitch and both visuospatial thickness and height. Infants looked significantly longer at cross-modally congruent stimuli in both experiments, suggesting a comparable starting point for height-pitch mappings and thickness-pitch mappings. This prelinguistic presence of space-pitch associations indicates that language is unlikely to create cross-modal mappings between space and pitch in the first place. Whereas language can influence space-pitch associations, it does not appear to establish these mappings. Instead, associations between space and pitch exist initially independently of language (in line with suggestions by Lakoff, 1993).

The nature of space-pitch representations

To explore the nature of space-pitch representations, in chapter 4 I investigated the locus of height-pitch associations, testing Lakoff's claim that metaphorical mappings are not 'in' language (i.e., they are not inherently linguistic). Whereas it is common in English to use the words *high* and *low* to talk about pitch, it is difficult to use the words *tall* and *short* sensibly to refer to pitch.

Still, both high/low and tall/short refer to the poles of a vertical spatial continuum. It was reasoned that if conventional space-pitch metaphors in language function by activating an association between a non-linguistic representation of pitch and a particular spatial schema (i.e., for English speakers, a vertical schema), then spatial terms that activate the same spatial schema should function similarly whether or not these terms can be used conventionally to talk about pitch. In speeded classification tasks, we found that participants responded faster to congruent compared to incongruent trials (congruency effect) when asked to classify pitches using the conventional words 'high' and 'low', and also when using unconventional words 'tall' and 'short'. This finding cannot be attributed to markedness or polarity alignment, since a control experiment showed no congruency effect for another pair of spatial terms (front/back) which also name the poles of a marked spatial continuum, but which activate the wrong spatial schema (sagittal instead of vertical). Words activating a vertical schema are thus sufficient to trigger space-pitch associations even if the words are lexically inappropriate, supporting the proposal that space-pitch metaphors are not inherently linguistic; rather, linguistic metaphors (whether conventional or novel) activate a mental metaphor by which English speakers conceptualize pitch, in part, in terms of vertical space.

To further investigate the nature of space-pitch representations, I explored spatial representations that may underlie musical pitch processing by using functional Magnetic Resonance Imaging (fMRI). Although numerous behavioral experiments confirm links between space and musical pitch, they are not informative about the neuronal underpinnings of space-pitch associations. To what extent do pitch representations activate neural circuitry that is involved in spatial processing and what is the exact nature of the spatial representations underlying musical pitch? In chapter 5 I directly compared different input modalities of spatial height (vision and touch) to musical pitch judgments. Participants completed visual and tactile height judgment tasks as well as pitch comparisons while lying in the MRI scanner. I reasoned that if pitch judgments rely on the same neural circuitry as judgments concerning visual, tactile, or multimodal space (i.e., visual *and* tactile space), one should see pitch activations in region(s) related to spatial height processing. Indeed, judgments of musical

pitch at least partially activated *visual* areas involved in spatial height processing. This finding suggests that space-pitch associations may involve modality-specific spatial representations, confirming a core assumption of embodied theories of metaphor.

Open questions and puzzles

What levels of pitch representation does language influence?

We show that language can influence pitch representations, but some questions remain open: Does using space-pitch metaphors in language influence how people *perceive* pitches, or only how they represent them, post-perceptually? The question of whether language influences perception, *per se*, has been debated extensively in linguistic relativity research (e.g., Boroditsky, Schmidt, & Phillips, 2003; Clifford et al., 2012; Casasanto et al., 2004; Gilbert, Regier, Kay, & Ivry, 2005; Lupyan, 2012; Roberson, Hanley, & Pak, 2009; see also Mitterer, Horschig, Müsseler, & Majid, 2009). In the color domain, for instance, there is evidence that linguistic categories can lead to differences in the visual mismatch negativity (vMMN), an index of automatic and preattentive change detection, suggesting that language effects may actually alter perceptual processing (Thierry, Athanasopoulos, Wiggett, Dering, & Kuipers, 2009). Although chapter 2 demonstrates an impact of language on pitch reproduction, our results do not elucidate the exact locus of this effect. Space-pitch interference effects could occur at any level of processing, including attention, perception, memory and/or tone production. In order to more precisely determine the level of language-specific effects, other experimental paradigms are needed. A crossmodal temporal order judgment (TOJ) task, for instance, could provide a fruitful way to test whether language influences space-pitch associations at a more perceptual level (Parise & Spence, 2009). In Parise's and Spence's study, participants had to judge whether a brief auditory or visual stimulus had been presented second (temporal order judgment, TOJ). The visual stimulus consisted of a visual circle that was either small or big. The auditory stimulus consisted of a tone that was either low (300Hz) or high (4500 Hz) in frequency. The auditory and visual stimuli presented on each trial were either crossmodally congruent

(e.g., big circle & low pitch) or incongruent (e.g., big circle & high pitch). Parise and Spence (2009) reasoned that if crossmodal associations really modulate audiovisual integration at a perceptual level, then participants should find it more difficult to indicate which stimulus had been presented second on crossmodally congruent as compared to crossmodally incongruent trials. In line with their hypothesis, participants found it harder to correctly resolve the temporal order for crossmodally congruent pairs of stimuli than for incongruent pairs. To find out whether language can have an impact on space-pitch integration at a perceptual level, Dutch and Farsi speakers (or speakers of other 'height' versus 'thickness'-languages) could be tested in a height and a thickness version of the temporal order judgment (TOJ) task. If language affects low-level audiovisual integration, one would expect to see contrasting patterns of temporal order judgments. Dutch speakers' TOJs should be affected more by crossmodally congruent height-pitch stimuli whereas Farsi speakers' TOJs should be affected more by crossmodally congruent thickness-pitch stimuli. On the other hand, however, it is possible that language-specific differences disappear at this level of perceptual integration.

Task-dependency and comparability of space-pitch associations

The preceding section already hinted at the possibility that varying experimental paradigms may lead to different outcomes with respect to space-pitch associations. Indeed we are faced with some seemingly contradictory evidence regarding space-pitch mappings. Whereas my results show that English speakers represent pitch in terms of height but not thickness (and thus size), other studies demonstrate that English speakers also associate size with pitch when tested in different kinds of judgment tasks (e.g., Evans & Treisman, 2010; Parise & Spence, 2009). Moreover, when explicitly asked to map spatial thickness and pitch, speakers of German (a 'height' language) matched the stimuli in accordance with Farsi speakers' mappings (Shayan, Ozturk, Bowerman, & Majid, submitted). Likewise, while Farsi speakers seem to represent pitch in terms of thickness but not height, a nonlinguistic height-pitch mapping was also demonstrated in members of a remote Kreung hill tribe in northeastern

Cambodia who apparently do not use height-pitch metaphors in their language (Parkinson, Kohler, Sievers, & Wheatley, 2012).

These contradictory results may suggest that the linguistic impact of space-pitch associations is to some extent task-dependent (i.e. results may differ depending on the level at which space-pitch associations are measured e.g., perceptual integration versus linguistic judgments).

However, whereas this conclusion may be valid based on the observations and arguments provided earlier, it has to be noted that results of the different tasks may not necessarily be comparable. For instance, the finding that height-pitch mappings are present even when speakers do not use height-pitch metaphors in language (Parkinson et al., 2012) does not speak to the question of whether different linguistic metaphors can affect nonlinguistic pitch cognition. Therefore, in order to test potential differences between languages (and also to test for task-specific effects), a direct comparison of different language groups is necessary. Performance of at least two groups has to be compared across the same sets of tasks in order to detect language-specific effects (as was done in chapter 2 of this dissertation). Moreover, it is important to ensure the validity of the experimental paradigm that is used. If interested in nonlinguistic space-pitch mappings, for instance, one should make sure that the employed task indeed measures nonlinguistic cognition (unlike tasks that allow verbal encoding of the stimuli as is the case in most psychophysical studies, e.g., in Evans & Treisman, 2010).

In order to delineate language-specific and task-specific effects in space-pitch mappings, we need more studies that compare the same tasks across different populations (e.g., language groups) as well as different tasks within the same population.

Task-dependency seems to be even more problematic in studies with infants and children. Developmental changes pose additional challenges to the comparability of experimental paradigms: Is task 'A' in a particular age-group the same or at least comparable to task 'A' in a slightly older age-group? What then about different tasks across different age-groups? In chapter 3 I find that prelinguistic infants at the age of 4 months are sensitive to thickness-pitch (or size-pitch) mappings in a preferential looking task. Ten-month-olds, however,

tested in a different experimental set-up (a habituation-dishabituation paradigm), are not sensitive to size-pitch associations (Haryu & Kajikawa, 2012). Again, in a 2-forced choice association task, 2- to 3-year-old preschoolers were capable of making size-pitch associations (Mondloch & Maurer, 2004), whereas children under 13 years of age were unable to systematically map size to pitch in a more abstract association task (Marks, et al., 1987). While it is possible that these varying outcomes result from nonlinear developmental trajectories, it is more plausible that results diverge because of differences in the employed tasks (and their demands). Thus, in order to ensure more valid results regarding cross-modal mappings in infants and children, the same age group should be tested in different experimental paradigms to rule out patterns of task-dependency.

Asymmetry in space-pitch metaphors

One essential claim of metaphor theory entails that metaphoric relationships are asymmetric in nature (Lakoff & Johnson, 1980). That is, people encode target domains in terms of source domains, rather than the reverse. For instance, people mostly use spatial terms to talk about valence or time (e.g. He is feeling *down*; a *long* vacation) but they less often talk about space in terms of valence or time. The same asymmetry seems to hold true for spatial metaphors of musical pitch.

While evidence for an asymmetry in linguistic metaphors can be derived from different observations (e.g., frequency or diachronic analyses, see Ullman, 1963; Williams, 1976), one can ask whether these linguistic asymmetries also have a counterpart in nonlinguistic (metaphoric) cognition. Do target domains like valence, time, or pitch asymmetrically rely on space and thus necessarily involve representations of the source domain? Indeed, there is some evidence supporting asymmetric metaphoric representations in cognition (see e.g. Meier & Robinson, 2004, for valence; Casasanto & Boroditsky, 2008, for time; and Casasanto, 2010, for pitch). For instance, in space-pitch interference tasks, Casasanto (2010) found that spatial information interfered more with people's pitch representations than did pitch information with their spatial representations. Similar findings have been obtained in a number of psychophysical tasks (e.g., Evans & Treisman, 2010; Hubbard & Courtney, 2010).

On a critical note, however, some asymmetry effects that are attributed to metaphoric representations may stem from asymmetries in measurement (if different measures are used for source and target domains) or in processing (e.g. more general visual dominance) (see also Crawford, 2009, for a critical overview). Moreover, it has to be acknowledged that asymmetry should not be equated with unidirectionality (for a discussion, see Casasanto & Boroditsky, 2008). While the influence of source (e.g. space) on target (e.g. pitch) is supposed to be greater than the reverse, there can also be some influence of target on source domains (see e.g. Maeda, Kanai, & Shimojo, 2005 for influence of pitch on space). Despite these restrictions, there is some evidence in favor of asymmetries in metaphoric language and thought. However, it is yet to be determined whether asymmetries in language affect asymmetries in thought or whether the reverse is true.

Crossing the senses

Throughout my thesis, I tested space-pitch associations by using cross-modal paradigms. Alongside auditory pitches, spatial stimuli were presented in the visual (chapters 2, 3, 5) or in the tactile modality (chapter 5). But to what extent do space-pitch mappings have to be cross-modal (i.e. cross-sensory) in nature and how do these mappings relate to synesthetic associations? With regard to the latter question, it was argued that cross-modal associations such as space-pitch mappings may constitute a weak form of synesthesia (e.g., Martino & Marks, 2001). Support for this claim comes from the similarities of associations reported by synesthetes as well as non-synesthetes. For instance, both groups seem to associate higher pitches with images that are sharper, brighter, smaller and higher in space (e.g., Fernay, Reby, & Ward, 2012; Marks, Hammeal, & Bornstein, 1987; Ward, Huckstep, & Tsakanikos, 2006). However, whereas cross-modal associations and actual synesthetic mappings may bear some similarities, there are also important differences (e.g., Spence, 2011). Unlike many types of synesthesia which tend to be absolute in nature (e.g., a synesthetic person consistently associates a particular letter with a particular color; see Van Leeuwen, Petersson, & Hagoort, 2010), cross-modal associations seem to be of a more relative or relational kind (e.g., Marks, 2011; Pedley & Harper, 1959;

Spence, 2011). Thus, a tone of a particular frequency is not directly linked to a particular position in space. Rather the association depends on the context. A tone that is *relatively* high in pitch may be linked to a position that is *relatively* high in space. However, once a second tone is presented that is even higher in frequency, the space-pitch relation will shift since the newly presented tone is now associated with the high position (see Pedley & Harper, 1959, for experimental evidence).

Space-pitch mappings may thus differ from synesthetic associations. Moreover, space-pitch mappings are not necessarily *cross-modal*. That is, the "input" modality of space does not have to be visual or tactile (or any other modality) but space could also be manipulated within the auditory domain itself. In line with this claim, height-pitch associations have been found to be present when space is exclusively manipulated in the auditory modality (i.e. the vertical position of loudspeakers; Pratt, 1930; Sonnadara, Gonzalez, Hansen, Elliott, & Lyons, 2009). In chapter 4, we obtained Stroop-like congruity effects when participants classified pitches by responding to visually presented words (e.g., high and low). Yet, presumably the visual format of the words was not necessary: similar congruity effects would be expected if the words were presented auditorily, or if participants had to speak the spatial words aloud. Moreover, height-pitch correspondences have been demonstrated in congenitally blind participants, suggesting that associations between height and pitch do not necessarily have to bridge vision and audition (Roffler & Butler, 1968; Walker, 1985, as cited in Walker, 1987). Instead of a cross-modal relationship, perhaps space and pitch should be considered to have a *cross-domain* (or *cross-dimensional*) relationship, with pitch and space each constituting a domain of knowledge. The domain (or dimension) of pitch is necessarily experienced in the auditory modality, but the domain of space can be experienced and represented multimodally (e.g., by vision, touch, audition, or some combination) or perhaps amodally (Pylyshyn, 1973). Further research is required to determine whether the cross-domain association of space and pitch differs when space is processed in one modality versus another.

Thought on language effects

Cross-domain mappings like associations between space and pitch have been found to exist, at least to some extent, independently of language (e.g., in prelinguistic infants). Is it therefore possible that the associations themselves shape linguistic metaphors (in accordance with Lakoff's proposal; Lakoff, 1993)? Indeed, the height-pitch metaphor can be found in many languages such as Spanish, German and Polish (Rusconi et al., 2006), as well as non-Indo-European languages, like Japanese and Chinese. In all of these languages, 'high' refers to high frequency sounds and 'low' to low frequency sounds, thus following prelinguistic preferences. However, there are exceptions to this trend. Speakers of the Austronesian language 'Are'are for instance talk about pitch in a 'reversed' way. The terms *siho* 'to go down' or *hi hu'a* 'towards the bottom' are used to refer to a high-pitched tone, whereas *hane* 'to go up' or *hi uuru* 'towards the top' are used to refer to a low-pitched tone (Zemp & Malkus, 1979). Prelinguistic associations are therefore unlikely to be the only source of influence that constrains the way pitch gets lexicalized across languages. Metaphors in language may be susceptible to correlations of properties in the real world as well as other types of experience (e.g., musical training). Accordingly, it has been suggested that the reversed height-pitch metaphor of the 'Are'are people has emerged from an extensive tradition in panpipe music, which plays an important role in the culture (Zemp & Malkus, 1979). Whereas players have to move 'up' to reach the longer pipes that produce lower-pitched tones, the shorter pipes require 'downward' movement in order to produce higher-pitched tones. However, nonlinguistic associations do not necessarily have to find their way into language. Piano players, for instance, have been found to represent pitch in terms of horizontal space, with lower pitches represented on the left and higher pitches represented on the right side of the body (Lidji et al., 2007). Still, there is no conventional linguistic metaphor referring to 'left' vs. 'right' pitches. While nonlinguistic associations and experience may in some cases affect linguistic structure, the extent of this influence as well as its timeline are hard to determine.

Acquisition of linguistic space-pitch metaphors

Since infants already have space-pitch associations available prior to language, what does this mean for language acquisition? Whereas one could assume that the early presence of mappings also facilitates the acquisition of linguistic space-pitch metaphors, the opposite seems to be the case. There is evidence that at least height-pitch metaphors are acquired quite late during language development (e.g., Costa-Giomi & Descombes, 1996). Three- to 6-year-old children who are able to discriminate changes in pitch direction were found to still have difficulties in producing correct height-pitch metaphors (Webster & Schlenrich, 1982; Costa-Giomi & Descombes, 1996, however see Stalinski, Schellenberg, & Trehub, 2008). Similarly, when tested in a comprehension task, the majority of 5-year-old Dutch children do not appear to understand pitch metaphors (Dolscheid, Hunnius, Casasanto, & Majid, in preparation). The relatively late acquisition of linguistic pitch metaphors might stem from the fact that children conflate distinct attributes of sounds. Evidence in support of this is children's use of blended terminology, e.g., *high* to refer to loudness and pitch (Hair, 1981). More specifically, however, metaphors seem to pose additional demands in language acquisition. Pitch metaphors are inherently polysemous. Children have to acquire both spatial and sound meanings. This process is likely to be more complex than when a single meaning has to be acquired (see e.g., Johnson, 1992). In line with this proposal French speaking children were better able to label pitch stimuli when they were trained to describe tones using the single-meaning terms *aigu* and *grave* (a pair of antonyms used only to label high and low pitches) than the polysemous words *haut* and *bas* (which are used to refer to pitch and space) (Costa-Giomi & Descombes, 1996). Moreover, the mastery of both spatial and pitch meanings does not seem to take place simultaneously. Rather, metaphoric expressions seem to be acquired in an asymmetric way as was famously suggested by Clark (1973). Clark argued that when a child acquires space-time metaphors: "spatial expressions should appear before time expressions, and in particular, each term that can be used both spatially and temporally should be acquired in its spatial sense first" (p. 57). Linguistic space-pitch metaphors seem to follow the suggested order of acquisition. Whereas Dutch children reliably comprehend the spatial relations of

height, almost all children are unable to understand height-pitch metaphors (Dolscheid et al., in preparation).

Trajectories of pitch metaphor acquisition are likely to have implications for language and thought relationships during developmental time. As was suggested earlier, language may affect nonlinguistic pitch representations via mechanisms of associative learning. Over time, speakers of a 'height' language like Dutch may strengthen the height-pitch mapping at the expense of the thickness-pitch mapping – and vice versa for speakers of a 'thickness' language like Farsi. Whereas evidence in support of this associative learning account is provided by linguistic training experiments in adults (chapter 2), it is unclear how learning effects may take place during language acquisition. If language can indeed strengthen particular mappings due to repetitive use of metaphors, when does this learning process begin? Since metaphorical extensions of space-pitch terminology occur rather late in development, it is likely that effects of language on nonlinguistic mappings also only occur in late childhood, after children learn to use space-pitch expressions like adults (see Lucy & Gaskins, 2001 for other language-on-thought effects during late childhood). The time course over which children acquire space-pitch metaphors in language, and the precise learning mechanisms that give rise to cross-linguistic differences in pitch representations remain topics for future research.

Conclusions

In summary, my findings suggest a particular relationship between space-pitch metaphors in language and nonlinguistic space-pitch associations. Whereas language can influence nonlinguistic space-pitch mappings (chapter 2), space-pitch mappings do not appear to be inherently linguistic. Language does not affect space-pitch associations in an online manner (chapter 2), nor does it seem to establish space-pitch mappings in the first place (chapter 3). Moreover, even lexically inappropriate spatial words can activate space-pitch associations, so long as the 'wrong' words cue the right spatial schema (chapter 4). Links between space and pitch also appear to be instantiated, in part, outside of the brain's core language system, in visuo-spatial areas (chapter 5).

Thus, language can influence space-pitch mappings even though these mappings are not created by language, nor are they inherent in language. Presciently, Benjamin Lee Whorf made a similar suggestion more than half a century ago. He wrote that "[o]ur metaphorical system, by naming nonspatial experiences after spatial ones, imputes to sounds [...] qualities like the [...] motions of spatial experience", yet he also acknowledged that metaphorical thinking "undoubtedly arises from a deeper source" (Whorf, 1956, p. 155).⁶ Here I find empirical evidence supporting both the "deeper source" of space-pitch mappings in prelinguistic infants and the power of metaphorical language to shape these mappings.

⁶ Note that Whorf considers what he calls metaphorical thinking, suggestion, or synesthesia to be the same (see Whorf, 1956).

References

- Amorapanth, P. X., Widick, P., & Chatterjee, A. (2010). The neural basis for spatial relations. *Journal of Cognitive Neuroscience*, 22(8), 1739–1753.
- Ashley, R. (2004). Musical pitch space across modalities: Spatial and other mappings through language and culture. In P. W. . Lipscomb, R. Ashley, & R. Gjerdingen (Eds.), *Proceedings of the 8th International Conference on Music Perception and Cognition*. Adelaide: Causal Productions.
- Aziz-Zadeh, L., Wilson, S. M., Rizzolatti, G., & Iacoboni, M. (2006). Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Current Biology*, 16(18), 1818–1823.
- Baayen, H. R. (2009). *LanguageR*. R package.
- Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press.
- Barker, A. (2004). *Greek Musical Writings: Volume 2, Harmonic and Acoustic Theory* (Vol. 2). Cambridge University Press.
- 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.
- Bates, D., & Maechler, M. (2009). Package “lme4”(Version 0.999375-32): linear mixed-effects models using S4 classes. Available (April 2011) at <http://cran.r-project.org/web/packages/lme4/lme4.pdf>.
- Ben-Artzi, E., & Marks, L. E. (1995). Visual-auditory interaction in speeded classification: Role of stimulus difference. *Attention, Perception, & Psychophysics*, 57(8), 1151–1162.
- 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), 663–672.
- Bizley, J. K., & Walker, K. M. . (2010). Sensitivity and selectivity of neurons in auditory cortex to the pitch, timbre, and location of sounds. *The Neuroscientist*, 16(4), 453–469.
- Bloom, P. (2002). *How children learn the meanings of words*. Cambridge, MA: MIT press.
- Bloom, P., & Keil, F. C. (2001). Thinking through language. *Mind & language*, 16(4), 351–367.
- Boroditsky, L. (2000). Metaphoric structuring: Understanding time through spatial metaphors. *Cognition*, 75(1), 1–28.
- Boroditsky, L., Fuhrman, O., & McCormick, K. (2011). Do English and Mandarin speakers think about time differently? *Cognition*, 118(1), 123–129.

- Boroditsky, L., Schmidt, L. A., Phillips, W., Gentner, D., & Goldin-Meadow, S. (2003). Sex, syntax, and semantics. *Language in mind: Advances in the study of language and cognition* (pp. 61–80). Cambridge, MA: MIT Press.
- 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.
- Boulenger, V., Hauk, O., & Pulvermüller, F. (2009). Grasping ideas with the motor system: semantic somatotopy in idiom comprehension. *Cerebral Cortex*, 19(8), 1905–1914.
- Bowerman, M., & Choi, S. (2003). Space under construction: Language-specific spatial categorization in first language acquisition. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 387–427). Cambridge: MIT Press.
- Brandt, T., & Dieterich, M. (2006). The vestibular cortex: its locations, functions, and disorders. *Annals of the New York Academy of Sciences*, 871(1), 293–312.
- Casasanto, D. (2008). Who's afraid of the big bad Whorf? Crosslinguistic differences in temporal language and thought. *Language learning*, 58(1), 63–79.
- Casasanto, D. (2009). Embodiment of abstract concepts: good and bad in right- and left-handers. *Journal of Experimental Psychology: General*, 138(3), 351–367.
- Casasanto, D. (2010). Space for thinking. In V. Evans & P. Chilton (Eds.), *Language, cognition and space: The state of the art and new directions* (pp. 453–478). London: Equinox Publishing.
- Casasanto, D., Boroditsky, L., Phillips, W., Greene, J., Goswami, S., Bocanegra-Thiel, S., Santiago-Diaz, I., et al. (2004). How deep are effects of language on thought? Time estimation in speakers of English, Indonesian, Greek, and Spanish. Proceedings of the 26th Annual Meeting of the Cognitive Science Society (pp. 186–191). Presented at the 26th Annual Meeting of the Cognitive Science Society, Chicago: Lawrence Erlbaum.
- Casasanto, D., & Bottini, R. (2010). Can mirror-reading reverse the flow of time? In N. S. Hölscher & M. Shipley (Eds.), *Spatial Cognition VII* (pp. 335–345). Berlin Heidelberg: Springer.
- Casasanto, D., & Chrysikou, E. G. (2011). When Left Is “Right” Motor Fluency Shapes Abstract Concepts. *Psychological science*, 22(4), 419–422.
- Chen, J. Y. (2007). Do Chinese and English speakers think about time differently? Failure of replicating Boroditsky (2001). *Cognition*, 104(2), 427–436.
- Clark, H. H. (1973). Space, time, semantics, and the child. In T. E. Moore (Ed.), *Cognitive Development and the Acquisition of Language* (pp. 27–63). New York: Academic Press.

- Clark, H. H., Carpenter, P. A., & Just, M. A. (1973). On the meeting of semantics and perception. In W. G. Chase (Ed.), *Visual information processing* (pp. 311-382). Chicago: Academic Press.
- Clifford, A., Franklin, A., Holmes, A., Drivonikou, V. G., Özgen, E., & Davies, I. R. L. (2012). Neural correlates of acquired color category effects. *Brain and Cognition*, 80(1), 126-143.
- Cohen Kadosh, R., Brodsky, W., Levin, M., & Henik, A. (2008). Mental representation: What can pitch tell us about the distance effect. *Cortex*, 44, 470-77.
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological review*, 82(6), 407-428.
- Connell, L., Cai, Z. G., & Holler, J. (2012). Do You See What I'm Singing? Visuospatial Movement Biases Pitch Perception. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th Annual Meeting of the Cognitive Science Society*. Presented at the CogSci, Austin, TX: Cognitive Science Society.
- Corbetta, M., Miezin, F. M., Dobmeyer, S., & Shulman, & G. L., Petersen, S.E. (1990). Attentional modulation of neural processing of shape, color, and velocity in humans. *Science*, 248, 1556-1559.
- Corradi-Dell'Acqua, C., Hesse, M. D., Rumiati, R. I., & Fink, G. R. (2008). Where is a nose with respect to a foot? The left posterior parietal cortex processes spatial relationships among body parts. *Cerebral cortex*, 18(12), 2879-2890.
- Costa-Giomi, E., & Descombes, V. (1996). Pitch labels with single and multiple meanings: A study with French-speaking children. *Journal of Research in Music Education*, 44(3), 204-214.
- Crawford, L. E. (2009). Conceptual metaphors of affect. *Emotion Review*, 1(2), 129-139.
- Cupchik, G. C., Phillips, K., & Hill, D. S. (2001). Shared processes in spatial rotation and musical permutation. *Brain and cognition*, 46(3), 373-382.
- Cytowic, R. E., & Eagleman, D. M. (2009). *Wednesday is indigo blue: Discovering the brain of synesthesia*. Cambridge, Massachusetts: MIT Press.
- Davidoff, J., Davies, I., & Roberson, D. (1999). Colour categories in a stone-age tribe. *Nature*, 398(6724), 203-204.
- Day, R. S., & Wood, C. C. (1972). Interactions between linguistic and nonlinguistic processing. *Journal of the Acoustical Society of America*, 51(79), 247-252. (Abstract).
- Day, S. (1996). Synaesthesia and synaesthetic metaphors. *Psyche*, 2(32), 1-16.
- 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.

- Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General*, 122(3), 371.
- Dempsey, N. P. (2008). *The coordination of action: Non-verbal cooperation in jazz jam sessions*. PhD Thesis.
- Dingemanse, M. (2011). *The meaning and use of ideophones in Siwu*. PhD Thesis, Radboud University Nijmegen.
- Dirven, R., & Taylor, J. (1988). The Conceptualization of Vertical Space in English: The Case of Tall. In B. Rudzka-Ostyn (Ed.), *Topics in Cognitive Linguistics* (pp. 207–229). Amsterdam: Benjamins.
- Dolscheid, S. (2008). *Übereinzelsprachliche Tendenzen bei synästhetischen Metaphern*. Retrieved from <http://miami.uni-muenster.de/servlets/DocumentServlet?id=4635>
- 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). Presented at the CogSci 2012, Austin, TX: Cognitive Science Society.
- Dolscheid, S., Shayan, S., Majid, A., & Casasanto, D. (2013). The Thickness of Musical Pitch: Psychophysical evidence for linguistic relativity. *Psychological Science*, 24(5), 613–621.
- 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.
- Eitan, Z., & Timmers, R. (2010). Beethoven's last piano sonata and those who follow crocodiles: Cross-domain mappings of auditory pitch in a musical context. *Cognition*, 114(3), 405–422.
- Evans, K. K., & Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features. *Journal of vision*, 10(1), 6–12.
- Fernay, L., Reby, D., & Ward, J. (2012). Visualized voices: A case study of audio-visual synesthesia. *Neurocase*, 18(1), 50–56.
- Fischer, M. H., Mills, R. A., & Shaki, S. (2010). How to cook a SNARC: number placement in text rapidly changes spatial-numerical associations. *Brain and cognition*, 72(3), 333–336.
- Foster, N. E. V., & Zatorre, R. J. (2010). Cortical structure predicts success in performing musical transformation judgments. *NeuroImage*, 53(1), 26–36.
- Friston, K. J., Holmes, A. P., Price, C. J., Büchel, C., & Worsley, K. J. (1999). Multisubject fMRI studies and conjunction analyses. *Neuroimage*, 10(4), 385–396.
- 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.

- Gallace, A., & Spence, C. (2006). Multisensory synesthetic interactions in the speeded classification of visual size. *Attention, Perception, & Psychophysics*, 68(7), 1191–1203.
- Garner, W. R. (1974). *The processing of information and structure*. Potomac: Lawrence Erlbaum.
- Gennari, S. P., Sloman, S. A., Malt, B. C., & Fitch, W. (2002). Motion events in language and cognition. *Cognition*, 83(1), 49–79.
- Gentner, D. (2002). Analogy in scientific discovery: The case of Johannes Kepler. In L. Magnani & N. J. Nersessian (Eds.), *Model-based reasoning: Science, technology, values* (pp. 21–39). New York: Kluwer Academic/ Plenum Publisher.
- Gentner, D. (2003). Why we're so smart. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 195–235). Cambridge, MA: MIT Press.
- Gentner, D., & Wolff, P. (2000). Metaphor and knowledge change. In E. Districh & A. Marbnau (Eds.), *Cognitive dynamics: Conceptual change in humans and machines* (pp. 295–342). Mahwah, NJ: Lawrence Erlbaum Associates.
- Ghazanfar, A. A., & Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends in cognitive sciences*, 10(6), 278–285.
- Gilbert, A. L., Regier, T., Kay, P., & Ivry, R. B. (2006). Whorf hypothesis is supported in the right visual field but not the left. *Proceedings of the National Academy of Sciences of the United States of America*, 103(2), 489–494.
- Gopnik, A., & Meltzoff, A. N. (1997). *Words, thoughts, and theories* (Vol. 1). MIT Press Cambridge, MA.
- Goy, A. (2002). Grounding meaning in visual knowledge. A case study: Dimensional adjectives. In R. K. Coventry & P. Olivier (Eds.), *Spatial Language: Cognitive and Computational Perspectives* (pp. 121–145). Dordrecht: Kluwer Academic Publishers.
- Hair, H. I. (1981). Verbal identification of music concepts. *Journal of Research in Music Education*, 29(1), 11–21.
- 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.
- Haspelmath, M. (2006). Against markedness (and what to replace it with). *Journal of Linguistics*, 42(1), 25–70.
- Haxby, J. V., Grady, C. L., Horwitz, B., Ungerleider, L. G., Mishkin, M., Carson, R. E., Herscovitch, P., et al. (1991). Dissociation of object and spatial visual processing pathways in human extrastriate cortex. *Proceedings of the National Academy of Sciences*, 88(5), 1621–1625.
- Heider, E. R., & Olivier, D. C. (1972). The structure of the color space in naming and memory for two languages. *Cognitive Psychology*, 3(2), 337–354.

- Hubbard, T. L. (1996). Synesthesia-like mappings of lightness, pitch, and melodic interval. *The American journal of psychology*, 219–238.
- Hubbard, T. L., & Courtney, J. R. (2010). Cross-modal influences on representational momentum and representational gravity. *Perception*, 39(6), 851–862.
- Huettel, S. A., Song, A. W., & McCarthy, G. (n.d.). *Functional magnetic resonance imaging*. 2004. Sunderland, Massachusetts: Sinauer.
- 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.
- Johnson, C. J. (1992). Cognitive components of naming in children: Effects of referential uncertainty and stimulus realism. *Journal of Experimental child psychology*, 53(1), 24–44.
- Koshino, H., Carpenter, P. A., Keller, T. A., & Just, M. A. (2005). Interactions between the dorsal and the ventral pathways in mental rotation: An fMRI study. *Cognitive, Affective, & Behavioral Neuroscience*, 5(1), 54–66.
- Lakens, D. (2012). Polarity correspondence in metaphor congruency effects: Structural overlap predicts categorization times for bipolar concepts presented in vertical space. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(3), 726–736.
- Lakoff, G. (1993). The contemporary theory of metaphor. In A. Ortony (Ed.), *Metaphor and thought* (Vol. 2, pp. 202–251). London/New York: Cambridge University Press.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by* (Vol. 111). Chicago London: The University of Chicago Press.
- Lakoff, G., & Johnson, M. (1999). *Philosophy in the Flesh: The embodied mind and its challenge to Western thought*. Basic books.
- Lancaster, J. L., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, C. S., Rainey, L., Kochunov, P. V., et al. (2000). Automated Talairach atlas labels for functional brain mapping. *Human brain mapping*, 10(3), 120–131.
- Landsberg, M. E. (1995). *Syntactic iconicity and linguistic freezes: the human dimension* (Vol. 9). De Gruyter Mouton.
- Lehrer, A. (1985). Markedness and antonymy. *Journal of Linguistics*, 21(2), 397–429.
- Levinson, S. C., & Majid, A. (29-31). The language of sound. In A. Majid (Ed.), *Field Manual Volume 10*. Nijmegen: Max Planck Institute for Psycholinguistics.
- Lewis, M. P. (2009). *Ethnologue: Languages of the world* (Vol. 9). SIL international Dallas, TX.
- Lewkowicz, D. J., & Turkewitz, G. (1980). Cross-modal equivalence in early infancy: Auditory–visual intensity matching. *Developmental psychology*, 16(6), 597–607.
- Li, P., Abarbanell, L., Gleitman, L., & Papafragou, A. (2011). Spatial reasoning in Tenejapan Mayans. *Cognition*, 120(1), 33–53.

- 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.
- Lucy, J. A., & Gaskins, S. (2001). Grammatical categories and the development of classification preferences: A comparative approach. In M. Bowerman & S. C. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 257-283). Cambridge: Cambridge University Press.
- Lupyan, G. (2012). Linguistically modulated perception and cognition: The label-feedback hypothesis. *Frontiers in Psychology*, 3.
- Lyons, J. (1977). *Semantics*. vol. 2. Cambridge University Press.
- 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.
- Maeda, F., Kanai, R., & Shimojo, S. (2004). Changing pitch induced visual motion illusion. *Current Biology*, 14(23), R990-R991.
- Majid, A., Bowerman, M., Kita, S., Haun, D., & Levinson, S. C. (2004). Can language restructure cognition? The case for space. *Trends in cognitive sciences*, 8(3), 108-114.
- Marks, L. E. (1982). Bright sneezes and dark coughs, loud sunlight and soft moonlight. *Journal of Experimental Psychology: Human Perception and Performance*, 8(2), 177-193.
- Marks, L. E. (1987). On cross-modal similarity: Auditory-visual interactions in speeded discrimination. *Journal of Experimental Psychology: Human Perception and Performance*, 13(3), 384-394.
- Marks, L. E. (1989). On cross-modal similarity: The perceptual structure of pitch, loudness, and brightness. *Journal of Experimental Psychology: Human Perception and Performance*, 15(3), 586-602.
- Marks, L. E. (2011). Synesthesia, Then and Now. *Intellectica*, 1(55), 47-80.
- Marks, L. E., Hammeal, R. J., & Bornstein, M. H. (1987). Perceiving similarity and comprehending metaphor. *Monographs of the Society for Research in Child Development*, 52(1), 1-102.
- Martino, G., & Marks, L. E. (1999). Perceptual and linguistic interactions in speeded classification: Tests of the semantic coding hypothesis. *Perception*, 28, 903-924.
- Martino, G., & Marks, L. E. (2001). Synesthesia: Strong and weak. *Current Directions in Psychological Science*, 10(2), 61-65.
- McGlone, M. S. (2007). What is the explanatory value of a conceptual metaphor? *Language & Communication*, 27(2), 109-126.
- 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. (1989). Dimensional interaction between color and pitch. *Journal of Experimental Psychology: Human Perception and Performance*, 15(1), 69–79.
- Melara, R. D., & Marks, L. E. (1990). Processes underlying dimensional interactions: Correspondences between linguistic and nonlinguistic dimensions. *Memory & Cognition*, 18(5), 477–495.
- Melara, R. D., & O'Brien, T. P. (1987). Interaction between synesthetically corresponding dimensions. *Journal of Experimental Psychology: General*, 116(4), 323–336.
- Merrill, J., Sammler, D., Bangert, M., Goldhahn, D., Lohmann, G., Turner, R., & Friederici, A. D. (2012). Perception of words and pitch patterns in song and speech. *Frontiers in Psychology*, 3.
- Merritt, D. J., Casasanto, D., & Brannon, E. M. (2010). Do monkeys think in metaphors? Representations of space and time in monkeys and humans. *Cognition*, 117(2), 191–202.
- Miller, G. A., & Johnson-Laird, P. N. (1976). *Language and perception*. Cambridge: Cambridge University Press.
- Miller, R. (1986). *The structure of singing: system and art in vocal technique*. Schirmer Books New York.
- Mitterer, H., Horschig, J. M., Müsseler, J., & Majid, A. (2009). The influence of memory on perception: It's not what things look like, it's what you call them. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(6), 1557–1562.
- Mondloch, C. J., & Maurer, D. (2004). Do small white balls squeak? Pitch-object correspondences in young children. *Cognitive, Affective, & Behavioral Neuroscience*, 4(2), 133–136.
- Mondor, T. A., Zatorre, R. J., & Terrio, N. A. (1998). Constraints on the selection of auditory information. *Journal of Experimental Psychology: Human Perception and Performance*, 24(1), 66–79.
- Moore, B. C. (1994). *An introduction to the psychology of hearing* (3rd ed., Vol. 4). London: Academic press.
- Moore, B. C. (2010). Aspects of Auditory Processing Related to Speech Perception. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The handbook of phonetic sciences* (pp. 454–488). Blackwell Publishing.
- Murphy, G. L. (1996). On metaphoric representation. *Cognition*, 60(2), 173–204.
- Murphy, G. L. (1997). Reasons to doubt the present evidence for metaphoric representation. *Cognition*, 62(1), 99–108.
- Näätänen, R., & Alho, K. (2004). Mechanisms of attention in audition as revealed by the event-related potentials of the brain. In M. I. Posner (Ed.), *Cognitive neuroscience of attention* (pp. 194–206). New York: Guilford Press.
- Nichols, T., Brett, M., Andersson, J., Wager, T., & Poline, J. B. (2005). Valid conjunction inference with the minimum statistic. *Neuroimage*, 25(3), 653–660.

- Noesselt, T., Hillyard, S. A., Woldorff, M. G., Schoenfeld, A., Hagner, T., Jäncke, L., Tempelmann, C., et al. (2002). Delayed striate cortical activation during spatial attention. *Neuron*, 35(3), 575–587.
- Nuerk, H. C., Iversen, W., & Willmes, K. (2004). Notational modulation of the SNARC and the MARC (linguistic markedness of response codes) effect. *Quarterly Journal of Experimental Psychology Section A*, 57(5), 835–863.
- Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition*, 108(1), 155–184.
- 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.
- Parkinson, C., Kohler, P. J., Sievers, B., & Wheatley, T. (2012). Associations between auditory pitch and visual elevation do not depend on language: Evidence from a remote population. *Perception*, 41(7), 854–861.
- 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.
- Pedley, P. E., & Harper, R. S. (1959). Pitch and the vertical localization of sound. *The American Journal of Psychology*, 72(3), 447–449.
- 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.
- Pinker, S. (1997). *How the mind works*. NY: Norton.
- Platel, H., Price, C., Baron, J. C., Wise, R., Lambert, J., Frackowiak, R. S., Lechevalier, B., et al. (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.
- Pratt, C. C. (1930). The spatial character of high and low tones. *Journal of Experimental Psychology*, 13(3), 278–285.
- Proctor, R. W., & Cho, Y. S. (2006). Polarity correspondence: A general principle for performance of speeded binary classification tasks. *Psychological Bulletin; Psychological Bulletin*, 132(3), 416–442.
- Pylyshyn, Z. W. (1973). What the mind’s eye tells the mind’s brain: A critique of mental imagery. *Psychological bulletin*, 80(1), 1–24.
- 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.
- Raposo, A., Moss, H. E., Stamatakis, E. A., & Tyler, L. K. (2009). Modulation of motor and premotor cortices by actions, action words and action sentences. *Neuropsychologia*, 47(2), 388–396.

- Regier, T., & Kay, P. (2009). Language, thought, and color: Whorf was half right. *Trends in cognitive sciences*, 13(10), 439–446.
- Renier, L. A., Anurova, I., De Volder, A. G., Carlson, S., VanMeter, J., & Rauschecker, J. P. (2009). Multisensory integration of sounds and vibrotactile stimuli in processing streams for “what” and “where.” *The Journal of Neuroscience*, 29(35), 10950–10960.
- 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.
- Roberson, D., & Davidoff, J. (2000). The categorical perception of colors and facial expressions: The effect of verbal interference. *Memory & Cognition*, 28(6), 977–986.
- Roberson, D., Davidoff, J., Davies, I. R. L., & Shapiro, L. R. (2005). Color categories: Evidence for the cultural relativity hypothesis. *Cognitive psychology*, 50(4), 378–411.
- Roberson, D., Hanley, J. R., & Pak, H. (2009). Thresholds for color discrimination in English and Korean speakers. *Cognition*, 112(3), 482–487.
- Roberson, D., Pak, H., & Hanley, J. R. (2008). Categorical perception of colour in the left and right visual field is verbally mediated: Evidence from Korean. *Cognition*, 107(2), 752–762.
- Roffler, S. K., & Butler, R. A. (1968). Factors that influence the localization of sound in the vertical plane. *The Journal of the Acoustical Society of America*, 43(6), 1255–1259.
- Roland, P. E., & Friberg, L. (1985). Localization of cortical areas activated by thinking. *Journal of Neurophysiology*, 53(5), 1219–1243.
- 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., Umiltà, C., & Butterworth, B. (2006). Spatial representation of pitch height: the SMARC effect. *Cognition*, 99(2), 113–129.
- Sano, H., & Jenkins, B. K. (1991). A Neural Network Model for Pitch Perception. In P. M. Todd & G. Loy (Eds.), *Music and connectionism* (pp. 42–53). MIT Press.
- 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.
- Schultz, J., & Lennert, T. (2009). BOLD signal in intraparietal sulcus covaries with magnitude of implicitly driven attention shifts. *Neuroimage*, 45(4), 1314–1328.

- Schwenzer, M., & Mathiak, K. (2011). Numeric aspects in pitch identification: an fMRI study. *BMC neuroscience*, 12(1), 1–9.
- Shayan, S., Ozge, O., Bowerman, M., & Majid, A. (submitted). Spatial metaphor in language can promote the development of cross-modal mappings in children.
- Shayan, S., Ozturk, O., & Sicoli, M. A. (2011). The thickness of pitch: Crossmodal metaphors in Farsi, Turkish, and Zapotec. *The Senses and Society*, 6(1), 96–105.
- Shepard, R. N. (1982). Structural representations of musical pitch. In D. Deutsch (Ed.), *The psychology of music* (pp. 343–390). New York: Academic Press.
- Slobin, D. (1996). From “thought and language” to “thinking for speaking”. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70–86). Cambridge: Cambridge University Press.
- Smith, L. B., & Sera, M. D. (1992). A developmental analysis of the polar structure of dimensions. *Cognitive Psychology*, 24(1), 99–142.
- Sonnadara, R. R., Gonzalez, D. A., Hansen, S., Elliott, D., & Lyons, J. L. (2009). Spatial Properties of Perceived Pitch. *Annals of the New York Academy of Sciences*, 1169(1), 503–507.
- Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, 73(4), 971–995.
- Stalinski, S. M., Schellenberg, E. G., & Trehub, S. E. (2008). Developmental changes in the perception of pitch contour: Distinguishing up from down. *The Journal of the Acoustical Society of America*, 124, 1759–1763.
- Stone, R. M. (1981). Toward a Kpelle conceptualization of music performance. *The Journal of American Folklore*, 94(372), 188–206.
- Straw, A. D. (2008). Vision egg: an open-source library for realtime visual stimulus generation. *Frontiers in neuroinformatics*, 2, 1–10.
- Stumpf, C. (2006). *Tonpsychologie* (Vol. 1). Leipzig: Hirzel.
- Taylor, J. R. (2002). Near synonyms as co-extensive categories: “high” and “tall” revisited. *Language Sciences*, 25(3), 263–284.
- Thierry, G., Athanasopoulos, P., Wiggett, A., Dering, B., & Kuipers, J. R. (2009). Unconscious effects of language-specific terminology on preattentive color perception. *Proceedings of the National Academy of Sciences*, 106(11), 4567–4570.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, 15(1), 273–289.
- Ueda, K., & Ohgushi, K. (1987). Perceptual components of pitch: Spatial representation using a multidimensional scaling technique. *The Journal of the Acoustical Society of America*, 82, 1193–1200.
- Ullmann, S. (1963). Semantic universals. In J. Greenberg (Ed.), *Universals of language*. (pp. 172–206). Cambridge: M.I.T Press.

- Ungerleider, L. G., & Haxby, J. V. (1994). "What" and "where" in the human brain. *Current opinion in neurobiology*, 4(2), 157–165.
- van Leeuwen, T. M., Petersson, K. M., & Hagoort, P. (2010). Synaesthetic colour in the brain: beyond colour areas. A functional magnetic resonance imaging study of synaesthetes and matched controls. *PloS one*, 5(8), e12074.
- Wagner, K., & Dobkins, K. R. (2011). Synaesthetic associations decrease during infancy. *Psychological science*, 22(8), 1067–1072.
- Wagner, S., Winner, E., Cicchetti, D., & Gardner, H. (1981). "Metaphorical" Mapping in Human Infants. *Child Development*, 728–731.
- Walker, B. N., & Ehrenstein, A. (2000). Pitch and pitch change interact in auditory displays. *Journal of Experimental Psychology: Applied*, 6(1), 15–30.
- Walker, K. M., Bizley, J. K., King, A. J., & Schnupp, J. W. (2011). Cortical encoding of pitch: recent results and open questions. *Hearing research*, 271(1), 74–87.
- 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.
- Walker, P., & Walker, L. (2012). Size–brightness correspondence: Crosstalk and congruity among dimensions of connotative meaning. *Attention, Perception, & Psychophysics*, 1–15.
- Walker, R. (1987). The effects of culture, environment, age, and musical training on choices of visual metaphors for sound. *Attention, Perception, & Psychophysics*, 42(5), 491–502.
- Ward, J., Huckstep, B., & Tsakanikos, E. (2006). Sound-colour synaesthesia: To what extent does it use cross-modal mechanisms common to us all? *Cortex*, 42(2), 264–280.
- Webster, P. R., & Schlentrich, K. (1982). Discrimination of pitch direction by preschool children with verbal and nonverbal tasks. *Journal of Research in Music Education*, 30(3), 151–161.
- Whorf, B. L. (1956). *Language, thought, and reality: Selected writings*. (J. B. Carroll, Ed.) (Vol. 5). MIT Press.
- Wiggett, A. J., & Davies, I. R. L. (2008). The effect of Stroop interference on the categorical perception of color. *Memory & cognition*, 36(2), 231–239.
- Willems, R. M., & Casasanto, D. (2011). Flexibility in embodied language understanding. *Frontiers in psychology*, 2.
- Williams, J. M. (1976). Synaesthetic adjectives: A possible law of semantic change. *Language*, 461–478.
- Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences*, 104(19), 7780–7785.
- Wojciulik, E., & Kanwisher, N. (1999). The generality of parietal involvement in visual attention. *Neuron*, 23(4), 747–764.

- Wong, P. C. M., Parsons, L. M., Martinez, M., & Diehl, R. L. (2004). The role of the insular cortex in pitch pattern perception: the effect of linguistic contexts. *The journal of neuroscience*, 24(41), 9153–9160.
- Wundt, W. (1893). *Logik*. [Eine Untersuchung der Prinzipien der Erkenntnis und der Methoden wissenschaftlicher Forschung] (Vol. 1). Stuttgart.
- Yu, N. (2003). Synesthetic metaphor: A cognitive perspective. *Journal of literary semantics*, 32(1), 19–34.
- Zangenehpour, S., & Zatorre, R. J. (2010). Crossmodal recruitment of primary visual cortex following brief exposure to bimodal audiovisual stimuli. *Neuropsychologia*, 48(2), 591–600.
- 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.
- Zatorre, R. J., Mondor, T. A., & Evans, A. C. (1999). Auditory attention to space and frequency activates similar cerebral systems. *Neuroimage*, 10(5), 544–554.
- Zemp, H., & Malkus, V. (1979). Aspects of 'Are'are musical theory. *Ethnomusicology*, 23(1), 5–48.

Samenvatting

Vaak worden ruimtelijke metaforen gebruikt om te verwijzen naar toonhoogte. In het Engels, Nederlands en vele andere talen kan de hoogte van een toon worden beschreven in ruimtelijke termen, zoals *hoog* of *laag* (bijv. Stumpf, 2006). Bovendien bestaat er psychofysisch bewijs dat toonhoogte en ruimtelijkheid met elkaar verbonden zijn in de hersenen (bijv. Melara & O' Brien, 1987; Evans & Treisman, 2010). Deze bevinding leidt tot een vraag die cruciaal is in mijn proefschrift: Wat is de relatie tussen ruimte-toonhoogte metaforen in taal en in het denken? In het bijzonder ben ik geïnteresseerd in de rol die taal speelt met betrekking tot associaties tussen toonhoogte en ruimtelijkheid. Weerspiegelt taal slechts niet-talige associaties of kan taal de verbanden tussen toonhoogte en ruimtelijkheid ook beïnvloeden? Welke rol spelen metaforen in talige associaties tussen ruimtelijkheid en toonhoogte in de hersenen? Wat zijn onderliggende neurale circuits van ruimte-toonhoogte metaforen?

Om deze vragen te beantwoorden combineer ik diverse methoden waaronder observaties in verschillende talen, psychofysische experimenten, studies met kinderen en neuroimaging.

Kan taal niet-talige toonhoogte-representaties beïnvloeden?

Om uit te zoeken of metaforen in taal van invloed kunnen zijn op ons denken heb ik gebruik gemaakt van cross-linguïstische diversiteit in het benoemen van toonhoogte. Waar sprekers van het Nederlands over *toonhoogte* praten (dus in termen van *hoogte*), gebruiken sprekers van het Farsi dikte-termen (dik = lage toon, dun = hoge toon). Deze talige verschillen kwamen ook naar voren bij het uitvoeren van twee niet-talige psychofysische imitatietaken. Toonhoogteschattingen van Nederlandse sprekers (in dit geval van gezongen reacties) werden significant beïnvloed door hoogte, maar niet door dikte. De toonhoogteschattingen van Farsi-sprekers daarentegen werden beïnvloed door dikte, maar niet door hoogte. Het lijkt er dus op dat sprekers van het Nederlands en het Farsi verschillende ruimtelijke representaties van toonhoogte hebben, overeenkomstig met de talige verschillen. Bovendien leken de prestaties van

sprekers van het Nederlands in de niet-talige toonhoogteproductietaak op die van Farsi-sprekers, nadat zij getraind waren in het gebruik van Farsi-achtige dikte-metaforen. Dit doet vermoeden dat taal een causale rol kan spelen in het vormen van verbanden tussen ruimtelijkheid en toonhoogte. In hoofdstuk 2 wordt aangetoond dat metaforen in taal impact kunnen hebben op niet-talige representaties van toonhoogte en ruimtelijkheid.

Maar *hoe* worden associaties tussen toonhoogte en ruimtelijkheid precies door taal beïnvloed? Om de aard van talige invloeden nader te onderzoeken werden de toon-ruimte associaties van sprekers van het Nederlands getest onder een verbale interferentie taak (Hoofdstuk 2). Anders dan bij andere effecten van ‘talige relativiteit’ (bijv. in het domein van kleuren; Winawer et al., 2007), bleef het effect van ruimte op toonhoogte zelfs bestaan toen proefpersonen bezig waren met het oefenen van lettersequenties. Daarom stel ik dat de resultaten van deze experimenten niet toegeschreven kunnen worden aan het online gebruik van verbale labels, maar eerder dat ze voortkomen uit analoge relaties tussen ruimte en toon in het lange termijn geheugen, die ten dele bepaald worden door taal.

De pre-linguïstische oorsprong van de verbanden tussen ruimte en toon(hoogte)

Taal weerspiegelt niet slechts bestaande ruimte-toon associaties, maar kan ook een effect hebben op niet-talige associaties. Mensen die andere ruimte-toonhoogte metaforen gebruiken in taal denken ook anders over toonhoogte. Hoewel talige metaforen zodoende dus meer lijken te zijn dan slechts ‘realisaties aan de oppervlakte’ van onderliggende verbanden (Lakoff, 1993) sluiten mijn resultaten niet uit dat ruimte-toonhoogte representaties zich misschien buiten het talige domein bevinden. Eerdere bevindingen die doen vermoeden dat hoogte-toon associaties misschien al bestaan voordat taal verworven is, ondersteunen dit voorstel (Walker et al., 2010; Wagner et al., 1981). Echter, veel over de oorsprong van andere ruimte-toon associaties is nog onbekend, zoals de relatie tussen dikte en toon. In hoofdstuk 3 heb ik daarom zowel ‘toondikte’ als ‘toonhoogte’ associaties bij pre-linguïstische kinderen getest. Mijn resultaten laten zien dat vier-maanden-oude baby’s gevoelig zijn voor verbanden tussen

auditiële toonhoogte en ruimtelijk visuele hoogte evenals dikte. In twee kiezend-kijken experimenten keken baby's significant langer naar stimuli die cross-modaal overeenstemden wat suggereert dat de beginpunten voor hoogte-toon en dikte-toon associaties gelijk zijn. De pre-linguïstische aanwezigheid van deze ruimte-toon associaties laat zien dat cross-modale verbanden tussen ruimte en toonhoogte waarschijnlijk geen talige oorsprong heeft. Hoewel taal ruimte-toonhoogte associaties weliswaar kan beïnvloeden, lijkt het niet waarschijnlijk dat taal deze associaties creëert. In plaats daarvan bestaan associaties tussen ruimte en toonhoogte aanvankelijk onafhankelijk van taal.

De aard van ruimte-toonhoogte representaties

Om de aard van ruimte-toonhoogte metaforen nader te onderzoeken heb ik in hoofdstuk 4 gekeken naar de lokalisering van hoogte-toon associaties, om Lakoff's claim dat metaforische mappings niet 'in' taal zitten (d.w.z. dat ze niet inherent talig zijn) te testen. Terwijl het in het Engels gebruikelijk is om de woorden *high* 'hoog' en *low* 'laag' te gebruiken om over toonhoogte te praten, is het moeilijk om de woorden *tall* 'lang' en *short* 'kort' daarvoor te gebruiken. Maar toch verwijzen zowel hoog/laag als lang/kort naar de uiteinden van een verticaal ruimtelijk continuüm. Onze redenering was dat als conventionele ruimte-toonhoogte metaforen in taal functioneren door het activeren van een associatie tussen niet-talige representaties van toonhoogte en een specifiek ruimtelijk schema (d.w.z. voor sprekers van het Engels een verticaal schema), ruimtelijke termen die dezelfde ruimtelijke schema's activeren dan hetzelfde zouden moeten functioneren om het even of de termen typisch gebruikt kunnen worden om over toonhoogte te praten. In versnelde classificatietaken hebben we ontdekt dat proefpersonen sneller reageerden op congruente items in vergelijking tot incongruente items (congruentie-effect) als hen gevraagd werd om tonen te classificeren waarbij ze de conventionele termen 'hoog' en 'laag', maar ook als ze de onconventionele termen 'kort' en 'lang' gebruikten. Deze bevinding kan niet worden toegeschreven aan gemarkeerdheid of polariteit, aangezien een controle-experiment geen congruentie-effect liet zien voor een derde paar ruimtelijke termen (voor/achter) die zich ook aan de uiteinden van een gemarkeerd ruimtelijk continuüm bevonden, maar die het verkeerde ruimtelijke

schema activeerden (sagittaal in plaats van verticaal). Woorden die een verticaal schema activeren volstaan dus om een ruimte-toonhoogte associatie te activeren, zelfs als de woorden lexicaal niet passend zijn. Dit ondersteunt het voorstel dat ruimte-toonhoogte metaforen niet inherent talig zijn; het is eerder zo dat talige metaforen (zowel conventionele of nieuwe) een mentale metafoor activeren waarbij sprekers van het Engels toonhoogte, ten dele, conceptualiseren in termen van verticale ruimtelijkheid.

Om de aard van ruimte-toonhoogte representaties nader te onderzoeken, heb ik de ruimtelijke representaties die ten grondslag liggen aan de verwerking van toonhoogte nader onderzocht met behulp van functional Magnetic Resonance Imaging (fMRI). Hoewel talrijke gedragsexperimenten het verband tussen ruimte en toonhoogte bevestigen, verschaffen ze geen informatie voor wat betreft de neurale onderbouwing van ruimte-toonhoogte associaties. In hoeverre activeren toonhoogterepresentaties de neurale circuits die bij ruimtelijke verwerking betrokken zijn en wat is de precieze aard van de ruimtelijke representaties die ten grondslag liggen aan toonhoogte? In hoofdstuk 5 heb ik input van verschillende (namelijk visuele en tactiele) modaliteiten met betrekking tot ruimtelijke hoogte en toonhoogte vergeleken. Proefpersonen deden visuele en tactiele hoogteschattingen alsmede vergelijkingen van toonhoogtes, terwijl ze in de MRI-scanner lagen. Ik redeneerde dat als toonhoogteschattingen dezelfde neurale circuits nodig hebben als schattingen aangaande visuele, tactiele of multimodale ruimtelijkheid (d.w.z visuele *en* tactiele ruimte), men toonhoogte-activering zou moeten zien in gebieden gerelateerd aan de verwerking van ruimtelijke hoogte. En inderdaad, schattingen van toonhoogte activeerden ten dele de *visuele* gebieden die betrokken zijn bij de verwerking van ruimtelijke hoogte. Deze bevinding doet vermoeden dat er modaliteit-specifieke ruimtelijke representaties betrokken zijn bij ruimte-toonhoogte associaties.

Conclusies

Samengevat laten mijn bevindingen zien dat er een bijzondere relatie is tussen ruimte-toonhoogte metaforen in taal en niet-talige ruimte-toonhoogte associaties. Hoewel taal niet-talige verbanden kan beïnvloeden (hoofdstuk 2),

lijken deze verbanden niet inherent talig. Taal heeft geen invloed op ruimte-toonhoogte associaties online (hoofdstuk 2) noch lijkt het er de oorsprong van te zijn (hoofdstuk 3). Bovendien kunnen zelfs lexicaal ongeschikte ruimtelijke woorden ruimte-toonhoogte associaties activeren zolang de 'foute' woorden maar aan het juiste ruimtelijke schema refereren (hoofdstuk 4). Verbanden tussen ruimte en toonhoogte lijken ook, ten dele, te bestaan buiten de kern van het taalsysteem, maar binnen de visueel ruimtelijke gebieden van de hersenen (hoofdstuk 5).

Op deze manier kan taal ruimte-toonhoogte verbanden beïnvloeden, ook al worden ze niet voortgebracht door taal en zijn ze ook niet inherent aan taal.

Acknowledgments

Writing a dissertation is a "journey", a "wave", or a "circus act" ... Looking at the acknowledgments of my fellow PhD-students suggests that the process of getting a PhD may be a metaphor in itself. But having written an entire thesis on metaphors, I would like to stop here.

So whatever the PhD process is like – be it a journey or a circus act – you cannot do it without the help of many many people! Thank you, bedankt, danke!

First of all, my gratitude goes to my advisors: Daniel Casasanto, Peter Hagoort, & Asifa Majid. Thank you for all your help and support!

Daniel, thanks for your incredibly helpful comments on my work and your advice throughout the years! You are so passionate about academia. Your enthusiasm is contagious and has kept me going even when I got lost at times.

Peter, thanks for giving me a home in the NBL department. Even though your schedule is tight, you were always there to discuss questions or problems, be it academic or non-academic ones. Thanks!

Last but certainly not least, I would like to thank you, Asifa. Your work and your ideas have been a source of inspiration all along. Thanks for sharing your knowledge, for mind-challenging brainstorm-sessions and that you invited me to be a member of the CATs project.

Thanks to my reading committee: Ton Dijkstra, Rolf Zwaan, & Andrea Bender. Thank you for taking the time to read and evaluate my thesis.

I would like to thank the International Max Planck Research School (IMPRS) for Language Sciences and the Max Planck Gesellschaft. My fellowship not only allowed me to pursue the topics that fascinate me but also to try out various kinds of methods. I am very grateful for this. Dirkje, Els, & Rachel – thanks for your help and guidance throughout my PhD years. Especially, thanks to you, Rachel, for your support at all times!

I would also like to thank all my colleagues and friends from the NBL department. It was challenging but great to learn so much about very different topics, all related to language and the brain. A special thanks to my fellow NBL-PhD-students for all your helpful comments during our weekly PhD meetings. Thanks, Roel, for teaching me how to conduct an fMRI study and analyze the data. Thanks to Paul Gaalman: Scanning subjects was fun because you were around! Thank you for your help, support, and your jokes about the Germans and their accent. Thanks Tildie and Ina, nothing would have worked without you! Ina, thanks for our nice chats and our dropjes-exchange. Katrien, thanks for being so gentle when I repeatedly stole (ehem borrowed) balistos from you and thanks for sharing the ups and downs of a PhD life. Thanks to the amazing RAs of the NBL-group. Especially, I would like to thank Manu, Thomas, Laura, & Marjolijn. Manu and Marjolijn, thanks for (re-)coding the baby data! Marjolijn, thanks for your help with testing kids and the archiving!

Also thank you to my colleagues and friends from the L&C-department, especially the CATs project. Thanks to Steve Levinson for inviting me to this project and for bringing the IMPRS for Language Sciences to life.

A special thanks to the BRC (Baby Research Center) team: Angela, Margret, & Sabine – thanks so much for your help! Angela, thanks for your help with recruiting, testing, and all the nice chats we've had! Sabine, vielen Dank für unsere tolle Zusammenarbeit! Danke, dass du immer ein offenes Ohr für alles hattest und für unsere inspirierenden Konversationen.

Thanks to all the parents and babies that participated in my studies.

I would also like to thank the TGs of the MPI and the Donders. Special thanks to Alex Dukers, Marek Tyc, Reiner Dirksmeyer, Ronald Fischer, & Ad Verbunt for helping me fight the countless problems along the way.

Thanks to the wonderful library-team!

Danke, Marie-Luise für deine Unterstützung.

Thanks to Angela Heuts & Nanjo for your support in all sorts of issues!

I wish I could thank Melissa Bowerman and Clemens Peter Herbermann in person. Melissa, thank you for your inspiring work, your kindness and your help during difficult moments. Herr Herbermann, danke, dass Sie mich erstmalig mit den synästhetischen Metaphern in Kontakt gebracht haben. Seitdem lässt mich das Thema nicht mehr los.

Vielen Dank auch an Pienie Zwitserlood und Nikolaus Himmelmann, die in mir während meines Studiums die Begeisterung für Sprache, Linguistik und Psycholinguistik geweckt haben.

I would like to thank the Relations-in-Relativity team: It was so great to organize this workshop with you – from the initial planning stages to its final showdown. Special thanks to you, Rebecca: You have made all this possible. Also thanks for our nice conversations, our tea-breaks and our amazing trip to Japan!

Thanks to Geoff, Roberto, Tom, & Luise for hosting me in New York! Thanks to you I had a wonderful time over there.

Furthermore thanks to all the people who participated in my experiments: Thanks for singing, pressing buttons, or enduring the fMRI-scanning.

Vielen Dank an meinen Super-Golf, der mich getreu (oder zumindest zumeist) an alle Ziele gebracht hat.

My dear MPI-volleyball-team: What would I have done without you? On Thursdays I was able to forget everything and just enjoyed playing with you. No more worries about data, abstracts, posters and deadlines, just the ball and us. And: lots of laughter! Thanks for so many lovely matches! I miss you.

Thanks, Shakila! For a good collaboration and nice chats.

Thanks to my (MPI) friends:

Tyko: Dank dir konnte ich die graue Zeit in Kleve überstehen und hatte sogar Spaß! Danke für geteilte Reftermahlzeiten, Psychohygiene, Waldspaziergänge, Boten- und Fahrdienste und für deine Freundschaft seit mittlerweile vielen Jahren.

Inge and Helen: I am so grateful that we met on day 1 of my PhD! Thanks for always hosting me, cheering me up (especially in the beginning), and for sharing laughter, tears and chocolate.

Christina: Danke für Carpooling, Katastrophenhilfe und die vielen guten Gespräche über das Leben und die Jobs ;).

Salomi(ta): I am so happy that I met you. Thanks for all the wonderful food you shared with me and for your loving and kind personality!

Sho: Mein liebster fellow-Cashot, mit dir hat alles Spaß gemacht, das Dance-your-Phd-projekt, unsere Babylablachkrämpfe und sogar die Laptopverluste. Wie gut, dass wir uns in Nijmegen gefunden haben.

Francisco, thanks for nice walks, good conversations and teaching me how to ski! Rósa, Rebecca, Ewelina, Marijt, Nienke, Katja, Jiyoun, Wencui, Kaoru, Sebastian, Mark, Flo, Manu, Judith, Line, Francisco, Dan, Katie, & Eddie – thanks for making my PhD-life so pleasant.

Thanks to my dear paranympths Tyko and Inge: I am very happy that you will be standing next to me during the big day.

Thanks to my old and new colleagues and friends at Cologne University: Sonja, Volker, & Uta, schön, dass wir jetzt alle in Köln sind! Ruth, Alina, Esraa, Sandra, Martina, Markus, & Uwe, dank euch habe ich mich direkt wohl in meiner neuen Heimat gefühlt.

Ich danke meinen Freunden außerhalb des MPIs:

Hanni, alias Dr. Schröder, bester Freund auf allen Spazierwegen, danke, dass ich so viel mit dir teilen durfte.

Nicole, gerade ohne deine Hilfe am Anfang hätte ich wohl sehr schnell aufgegeben. Danke, dass du mich immer ermutigt hast und für unsere tiefe Freundschaft über die Kontinente hinweg!

Karo, danke für unsere schönen Reisen nach Taizé, die mir immer wieder Kraft gegeben haben.

Tabbi, danke, dass du so eine tolle Mitbewohnerin warst und man einfach auf dich zählen kann.

Kathrin, Käthe, Kajüte, danke, dass du alle Zweifel immer fortgefeht hast!

Jani, meine beste Freundin, ohne dich hätte ich das alles nicht durchgestanden! Danke.

Am meisten möchte ich meiner Familie danken: Danke, dass ihr immer für mich da wart und mich in allem unterstützt habt. Peter Dolscheid, Monika Dolscheid, Dr. Ramona Dolscheid ;), Raphael Dolscheid und Agnes Baake: Vielen Dank für moralische, kulinarische und finanzielle Unterstützung, Autopannenhilfe, Aufmunterung, Basteln von Experimentmaterialien und dass ihr immer als meine ersten Versuchspersonen zur Verfügung gestanden habt!

Heiko, ohne dich geht gar nichts. Danke für deine Unterstützung, deine Liebe und deinen Halt in all den Jahren. Dir sei dieses Buch gewidmet.

Curriculum Vitae

Sarah Dolscheid has studied General Linguistics, Communication Science, and Psychology at University of Münster and University of California, Los Angeles (UCLA). After graduating from Münster University, she was awarded a fellowship of the International Max Planck Research School (IMPRS) for Language Sciences at the Max Planck Institute for Psycholinguistics. For her doctoral thesis she conducted interdisciplinary research-projects at the MPI, the Donders Institute, the Baby Research Center Nijmegen and The New School for Social Research, New York. She currently works as a research fellow in the Department of Rehabilitation and Special Education at University of Cologne.

List of Publications

Dolscheid, S., Graver, C., & Casasanto, D. (2013). Spatial congruity effects reveal metaphors, not markedness. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 2213-2218). Austin, TX: Cognitive Science Society.

Dolscheid, S., Shayan, S., Majid, A., & Casasanto, D. (2013). The thickness of musical pitch: Psychophysical evidence for linguistic relativity. *Psychological Science*, 24(5), 613-621.

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.), *CogSci 2012: Proceedings of the 34th Annual Meeting of the Cognitive Science Society* (pp. 306-311). Austin, TX: Cognitive Science Society.

Dolscheid, S., Shayan, S., Majid, A., & Casasanto, D. (2011). The thickness of musical pitch: Psychophysical evidence for the Whorfian hypothesis. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 537-542). Austin, TX: Cognitive Science Society.

MPI Series in Psycholinguistics

1. The electrophysiology of speaking: Investigations on the time course of semantic, syntactic, and phonological processing. *Miranda van Turenhout*
2. The role of the syllable in speech production: Evidence from lexical statistics, metalinguistics, masked priming, and electromagnetic midsagittal articulography. *Niels O. Schiller*
3. Lexical access in the production of ellipsis and pronouns. *Bernadette M. Schmitt*
4. The open-/closed-class distinction in spoken-word recognition. *Alette Haveman*
5. The acquisition of phonetic categories in young infants: A self-organising artificial neural network approach. *Kay Behnke*
6. Gesture and speech production. *Jan-Peter de Ruiter*
7. Comparative intonational phonology: English and German. *Esther Grabe*
8. Finiteness in adult and child German. *Ingeborg Lasser*
9. Language input for word discovery. *Joost van de Weijer*
10. Inherent complement verbs revisited: Towards an understanding of argument structure in Ewe. *James Essegbey*
11. Producing past and plural inflections. *Dirk Janssen*
12. Valence and transitivity in Saliba: An Oceanic language of Papua New Guinea. *Anna Margetts*
13. From speech to words. *Arie van der Lugt*
14. Simple and complex verbs in Jaminjung: A study of event categorisation in an Australian language. *Eva Schultze-Berndt*
15. Interpreting indefinites: An experimental study of children's language comprehension. *Irene Krämer*
16. Language-specific listening: The case of phonetic sequences. *Andrea Weber*
17. Moving eyes and naming objects. *Femke van der Meulen*
18. Analogy in morphology: The selection of linking elements in Dutch compounds. *Andrea Krott*
19. Morphology in speech comprehension. *Kerstin Mauth*
20. Morphological families in the mental lexicon. *Nivja H. de Jong*
21. Fixed expressions and the production of idioms. *Simone A. Sprenger*
22. The grammatical coding of postural semantics in Goemai (a West Chadic language of Nigeria). *Birgit Hellwig*
23. Paradigmatic structures in morphological processing: Computational and cross-linguistic experimental studies. *Fermín Moscoso del Prado Martín*
24. Contextual influences on spoken-word processing: An electrophysiological approach. *Daniëlle van den Brink*
25. Perceptual relevance of prevoicing in Dutch. *Petra M. van Alphen*
26. Syllables in speech production: Effects of syllable preparation and syllable frequency. *Joana Cholin*
27. Producing complex spoken numerals for time and space. *Marjolein Meeuwissen*

28. Morphology in auditory lexical processing: Sensitivity to fine phonetic detail and insensitivity to suffix reduction. *Rachèl J. J. K. Kemps*
29. At the same time...: The expression of simultaneity in learner varieties. *Barbara Schmiedtová*
30. A grammar of Jalonke argument structure. *Friederike Lüpke*
31. Agrammatic comprehension: An electrophysiological approach. *Marlies Wassenaar*
32. The structure and use of shape-based noun classes in Miraña (North West Amazon). *Frank Seifart*
33. Prosodically-conditioned detail in the recognition of spoken words. *Anne Pier Salverda*
34. Phonetic and lexical processing in a second language. *Mirjam Broersma*
35. Retrieving semantic and syntactic word properties. *Oliver Müller*
36. Lexically-guided perceptual learning in speech processing. *Frank Eisner*
37. Sensitivity to detailed acoustic information in word recognition. *Keren B. Shatzman*
38. The relationship between spoken word production and comprehension. *Rebecca Özdemir*
39. Disfluency: Interrupting speech and gesture. *Mandana Seyfeddinipur*
40. The acquisition of phonological structure: Distinguishing contrastive from non-contrastive variation. *Christiane Dietrich*
41. Cognitive cladistics and the relativity of spatial cognition. *Daniel B.M. Haun*
42. The acquisition of auditory categories. *Martijn Goudbeek*
43. Affix reduction in spoken Dutch. *Mark Pluymaekers*
44. Continuous-speech segmentation at the beginning of language acquisition: Electrophysiological evidence. *Valesca Kooijman*
45. Space and iconicity in German Sign Language (DGS). *Pamela Perniss*
46. On the production of morphologically complex words with special attention to effects of frequency. *Heidrun Bien*
47. Crosslinguistic influence in first and second languages: Convergence in speech and gesture. *Amanda Brown*
48. The acquisition of verb compounding in Mandarin Chinese. *Jidong Chen*
49. Phoneme inventories and patterns of speech sound perception. *Anita Wagner*
50. Lexical processing of morphologically complex words: An information-theoretical perspective. *Victor Kuperman*
51. A grammar of Savosavo, a Papuan language of the Solomon Islands. *Claudia Wegener*
52. Prosodic structure in speech production and perception. *Claudia Kuzla*
53. The acquisition of finiteness by Turkish learners of German and Turkish learners of French: Investigating knowledge of forms and functions in production and comprehension. *Sarah Schimke*
54. Studies on intonation and information structure in child and adult German. *Laura de Ruiter*
55. Processing the fine temporal structure of spoken words. *Eva Reinisch*
56. Semantics and (ir)regular inflection in morphological processing. *Wieke Tabak*
57. Processing strongly reduced forms in casual speech. *Susanne Brouwer*

58. Ambiguous pronoun resolution in L1 and L2 German and Dutch. *Miriam Ellert*
59. Lexical interactions in non-native speech comprehension: Evidence from electroencephalography, eye-tracking, and functional magnetic resonance imaging. *Ian FitzPatrick*
60. Processing casual speech in native and non-native language. *Annelie Tuinman*
61. Split intransitivity in Rotokas, a Papuan language of Bougainville. *Stuart Robinson*
62. Evidentiality and intersubjectivity in Yurakaré: An interactional account. *Sonja Gipper*
63. The influence of information structure on language comprehension: A neurocognitive perspective. *Lin Wang*
64. The meaning and use of ideophones in Siwu. *Mark Dingemanse*
65. The role of acoustic detail and context in the comprehension of reduced pronunciation variants. *Marco van de Ven*
66. Speech reduction in spontaneous French and Spanish. *Francisco Torreira*
67. The relevance of early word recognition: Insights from the infant brain. *Caroline Junge*
68. Adjusting to different speakers: Extrinsic normalization in vowel perception. *Matthias J. Sjerps*
69. Structuring language. Contributions to the neurocognition of syntax. *Katrien R. Segaert*
70. Infants' appreciation of others' mental states in prelinguistic communication: A second person approach to mindreading. *Birgit Knudsen*
71. Gaze behavior in face-to-face interaction. *Federico Rossano*
72. Sign-spatiality in Kata Kolok: how a village sign language of Bali inscribes its signing space. *Conny de Vos*
73. Who is talking? Behavioural and neural evidence for norm-based coding in voice identity learning. *Attila Andics*
74. Lexical processing of foreign-accented speech: Rapid and flexible adaptation. *Marijt Witteman*
75. The use of deictic versus representational gestures in infancy. *Daniel Puccini*
76. Territories of knowledge in Japanese conversation. *Kaoru Hayano*
77. Family and neighbourhood relations in the mental lexicon: A cross-language perspective. *Kimberley Mulder*
78. Contributions of executive control to individual differences in word production. *Zeshu Shao*
79. Hearing speech and seeing speech: Perceptual adjustments in auditory-visual processing. *Patrick van der Zande*
80. High pitches and thick voices: The role of language in space-pitch associations, *Sarah Dolscheid*

