Research Article

Psychological Science
24(5) 613–621

DOI: 10.1177/0956797612457374
pss.sagepub.com

The Thickness of Musical Pitch: Psychophysical Evidence for Linguistic Relativity

Sarah Dolscheid[1,2], Shakila Shayan[1], Asifa Majid[1,3], and Daniel Casasanto[1,3,4]

[1]Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands; [2]International Max Planck Research School for Language Sciences; [3]Donders Center for Brain, Cognition, and Behaviour, Radboud University; and [4]Department of Psychology, The New School for Social Research

Abstract

Do people who speak different languages think differently, even when they are not using language? To find out, we used nonlinguistic psychophysical tasks to compare mental representations of musical pitch in native speakers of Dutch and Farsi. Dutch speakers describe pitches as high (*hoog*) or low (*laag*), whereas Farsi speakers describe pitches as thin (*nāzok*) or 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 influences mental representations of pitch, we trained native Dutch speakers to describe pitch in terms of thickness, as Farsi speakers do. After the training, Dutch speakers' performance on a nonlinguistic psychophysical task resembled the performance of native Farsi speakers. People who use different linguistic space-pitch metaphors also think about pitch differently. Language can play a causal role in shaping nonlinguistic representations of musical pitch.

Keywords

metaphor, musical pitch, psychophysics, space, Whorfian hypothesis, cross-cultural differences, language, music

Received 7/12/11; Revision accepted 7/7/12

In English, musical pitches can be called *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 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 press response keys that are spatially high more quickly in response to high-frequency pitches than in response to low-frequency pitches; the reverse is true for response keys that are spatially low (Lidji, Kolinsky, Lochy, & Morais, 2007; Rusconi, Kwan, Giordano, Umiltà, & Butterworth, 2006). Going beyond binary high-low correspondences, psychophysical pitch-reproduction tasks have shown 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 the minds of 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, and this representation is reflected 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*, respectively. 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

Corresponding Author:
Daniel Casasanto, The New School for Social Research, Department of Psychology, 80 Fifth Ave., 7th Floor, New York, NY 10011
E-mail: casasanto@alum.mit.edu

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, Bowerman, & Majid, 2012; Shayan, Ozturk, & Sicoli, 2011).

Do people who use different linguistic metaphors for pitch also mentally represent pitch differently? If so, how deep are the effects of language on mental representations of pitch? Does language shape the nonlinguistic representations that people use when perceiving or producing musical pitches, even when they are not using language? One hint that people who use different pitch metaphors think about pitch differently comes from co-speech gestures. For example, the Manza have been observed lowering their hand in space when referring to smaller (i.e., higher) pitches; this gesture shows a mental representation of pitch that is contrary to the high-low mapping seen in English speakers (Ashley, 2004). This suggests that people may conceptualize pitch in a way that is consistent with their pitch vocabulary. However, gestures that match the co-occurring speech may reveal conventions for communicating about musical pitches, rather than modes of conceptualizing them. Alternatively, such gestures may reveal a "shallow" influence of language on thought, indicating that people do indeed conceptualize pitch in language-specific ways, but only when 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. In the study reported here, we used a pair of psychophysical tasks with nonlinguistic stimuli and responses to test mental representations of pitch in speakers of a language with height metaphors for pitch (Dutch) and in speakers of a language with thickness metaphors for pitch (Farsi). In one task (height interference), participants saw lines at varying heights while they listened to tones of different pitches. After each tone ended, a microphone appeared on the screen, and participants reproduced the pitch by singing it. In the other task (thickness interference), participants saw lines of varying thickness while they heard tones of different pitches; they then sang the tones that they heard. In both tasks, the spatial information was irrelevant, and spatial variation was orthogonal to variation in pitch. Thus, 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 mental representations of pitch were similar in Dutch and Farsi speakers, irrespective of the languages they speak, performance on these tasks should not differ between these two language groups. However, if Dutch and Farsi speakers conceptualize pitch the way they talk about it, by activating different kinds of spatial representations, they should show contrasting patterns of cross-dimensional interference: Dutch speakers' pitch estimates should be more affected by irrelevant height information, and Farsi speakers' estimates should be more influenced by irrelevant thickness information.

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

Method

Participants. Native Dutch speakers (n = 20) and native Farsi speakers (n = 20) performed the height-interference task. Likewise, native Dutch (n = 20) and Farsi (n = 20) speakers performed the thickness-interference task. Within each language group, some participants were tested in both tasks, with the order of tasks counterbalanced and the testing sessions separated by at least 1 week. One additional Farsi-speaking participant was tested in the height-interference task but was excluded for not following the instructions. Dutch participants were recruited from the Max Planck Institute participant pool. Farsi speakers were recruited by one of the coauthors from Farsi-speaking communities in the Dutch cities of Nijmegen, Delft, and Leiden.[1] All participants received payment to participate.

Materials. For each trial of the height-interference task (adapted from Casasanto, 2010), participants viewed a computer screen showing a horizontal line intersecting a vertical reference line at one of nine different heights (the horizontal lines ranged from 80 to 720 pixels from the bottom of the screen in 80-pixel increments). For each trial of the thickness-interference task, a vertical line appeared in the middle of the computer 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 the height-interference task. Each line was presented simultaneously with a tone. In each task, the nine different 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 laptop PC (screen resolution = 1024 × 768 pixels) using Presentation software (Neurobehavioral Systems, Albany, CA). Lines were shown in white on a gray background (350 pixels wide). Auditory stimuli consisted of pure tones created using Audacity software (http://audacity.sourceforge.net/) and presented for 2 s through sealed headphones at a constant amplitude.

The participants' sung responses were recorded using an EDIROL R-09 recording device (Roland, Los Angeles, CA) and analyzed, using Praat software (Praat, Amsterdam, The Netherlands), by a coder blind to the 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 to 300 ms after the estimated temporal midpoint, to ensure that the measured F_0 was representative of the whole response.

Written instructions were presented in the native language of the participant. The instructions contained no space-pitch metaphors. The tasks themselves involved 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 the tones back as accurately as possible. After 3 practice trials, participants were presented with the 81 line-pitch pairings, one at a time, in random order, for 2 s each. Immediately after each stimulus, a picture of a microphone appeared in the center of the screen, indicating that the participant had 2 s to sing back the pitch. Each response period was followed by an intertrial interval of 500 ms. After 40 trials, participants had a self-paced break. Testing lasted about 15 min 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 (M = 2.65) against zero, $t(19) = 2.70$, $p = .01$ (all tests were two-tailed unless otherwise stated; for the ranges of averaged pitch estimates for all four of our experiments, see Table S1 in the Supplemental Material available online). Tones accompanied by higher lines were reproduced at a higher frequency, on average, than the same tones when accompanied by lower lines. By contrast, the thickness of the stimuli did not have a significant effect on pitch reproduction (mean slope = 0.60), $t(19) = 0.57$, n.s. Farsi speakers showed the opposite pattern: Thickness influenced their pitch estimates, with tones accompanied by thicker lines reproduced at a lower frequency, on average, than the same tones accompanied by thinner lines (mean slope = −2.85), $t(19) = 2.09$, $p = .05$. Height did not have a significant effect on pitch reproduction in Farsi speakers (mean slope = −0.71), $t(19) = 1.16$, n.s.

To test for the predicted interaction of language (Dutch, Farsi) and task (height interference, thickness interference), we multiplied the normalized slopes from the thickness-interference task by −1. This multiplication was necessary because the relationship between spatial magnitude and frequency is reversed between the 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 aligned the space and pitch continuums (i.e., the slopes then indicated the same relationship between spatial magnitude and frequency for the two tasks). According to a 2 × 2 analysis of variance (ANOVA), language interacted with task to predict the effect of space on pitch estimates, $F(1, 79) = 10.73$, MSE = 0.29, $p = .002$, a finding consistent with the use of height metaphors in Dutch and thickness metaphors in Farsi (Fig. 1a; see also Fig. S1 in the Supplemental Material). There were no main effects. Additional analyses ensured that the observed effects could not be explained by cross-cultural differences in musical experience or proficiency (for more about these additional analyses, see the Supplemental Material).

In planned pairwise comparisons, the effect of height interference was greater in Dutch speakers than in Farsi speakers, $t(38) = 2.90$, $p = .01$. 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 in Farsi speakers, the effect of thickness interference was greater than the effect of height interference, $t(38) = 2.38$, $p = .02$.

Pitch estimation: within-domain effects. We conducted further analyses 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 the participant's reproduced pitches. We then calculated the average of the slopes for each combination of language group and task, and tested whether these means differed from zero. The obtained values were as follows: Dutch speakers, height interference: mean slope = 1.02, $t(19) = 12.50$, $p = .0001$; Dutch speakers, thickness interference: mean slope = 0.87, $t(19) = 12.68$, $p = .0001$; Farsi speakers, height interference: mean slope = 0.61, $t(19) = 6.67$, $p = .0001$; Farsi speakers, thickness interference: mean slope = 0.58, $t(19) = 6.21$, $p = .0001$. According to a 2 × 2 ANOVA, language did not interact with task to predict the effect of actual pitch on estimated pitch, $F(1, 79) = 0.46$, MSE = 0.002, n.s. Overall, Dutch speakers' pitch estimates were more accurate than Farsi speakers' estimates, $F(1, 79) = 16.96$, MSE = 0.002, $p = .0001$, but this main effect of language on within-domain performance cannot explain the Language × Task interaction we found in the cross-domain analysis.

Finally, we conducted a three-way ANOVA combining the cross-domain and within-domain analyses, after normalizing all values of space and pitch for each participant. There was a three-way interaction of language (Dutch,

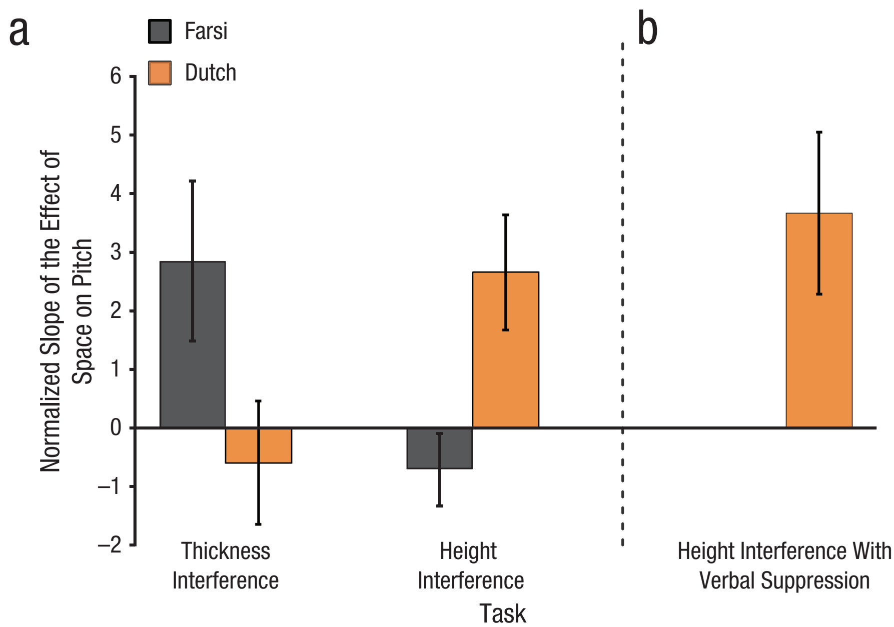

Fig. 1. Results of (a) Experiment 1 and (b) Experiment 2: normalized slope of the effect of the spatial stimulus on Farsi and Dutch speakers' pitch reproduction in the thickness- and height-interference tasks (Experiment 1) and the height-interference task with verbal suppression (Experiment 2). Error bars indicate standard errors of the mean.

Farsi), task (height interference, thickness interference), and domain (within-domain effects, cross-domain effects), $F(1, 159) = 7.04$, $MSE = 0.0001$, $p = .01$, indicating that the observed cross-dimensional interference effects cannot be explained by unpredicted differences in within-domain performance.

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

Experiment 2: What Are the Effects of 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? Although it is possible that participants attempted to label the stimuli, doing so could not have produced the observed effects. That is, covert labeling 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 using ordinary nontechnical words: Labels such as *high* and *low* are too coarse to have helped participants discriminate among nine randomly ordered pitches at the level of the semitone (e.g., the word *low* applied to all of the pitches on the low end of the continuum, and therefore could not help participants discriminate among them). Second, and more important, covertly labeling pitches as *high* or *low* would have worked against the observed spatial-interference effects because space and pitch varied orthogonally.

Still, to rule out the possibility of on-line language effects, we asked Dutch speakers to perform a version of the height-interference task that included verbal interference. If the effect of height on pitch in Experiment 1 was driven by covert labeling of the stimuli using spatial words, the effect should disappear under verbal interference. However, we hypothesized that this effect was caused not by on-line use of linguistic metaphors for pitch, but rather by 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 our hypothesis is correct, the effect of height on pitch in Dutch speakers should persist under conditions of verbal interference.

Method

Participants and materials. A new sample of native Dutch speakers ($N = 22$) participated for payment. The same materials as in the height-interference task of Experiment 1 were used. Additionally, 81 unpronounceable five-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 they had heard. Before each trial, they were shown one of the five-letter strings for 2 s. They had been instructed to rehearse the letters of each string silently. After one third of the trials, participants' recognition of the presented letter string was tested using a two-alternative forced-choice test. Participants responded by key press to indicate whether the string was the one presented at the start of the trial ("S" key) or a foil ("L" key). The next trial began after a 500-ms intertrial interval. After 40 trials, participants had a self-paced break. Testing lasted about 20 min.

Results and discussion

Verbal-interference task. Participants' recognition of the letter series was much greater than chance (mean accuracy = 85%, SD = 9.25), $t(21) = 17.76$, $p = .0001$. This finding indicates 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. Spatial height influenced pitch estimates under verbal interference (mean slope = 3.66), $t(21) = 2.65$, $p = .02$ (Fig. 1b; see also Fig. S2 in the Supplemental Material). The slopes for the Dutch speakers performing the height-interference task in Experiment 1 did not differ from the slopes for the Dutch speakers performing the same task under verbal interference in Experiment 2 (difference of normalized slopes = 1.01), $t(40) = 0.58$, n.s. Therefore, the effects of spatial stimuli on reproduction of pitch stimuli cannot be attributed to covert activation of verbal labels for the stimuli.[2]

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. We conducted a two-part training study to investigate whether language can play a causal role in pitch representation. In this experiment, Dutch speakers completed sentences about pitch relationships using Farsi-like thickness metaphors (thickness training) or using height metaphors that would be familiar to them (height training; control condition). To determine whether this linguistic training influenced nonlinguistic pitch representations, we then tested all participants using the thickness-interference task from Experiment 1. If experience using linguistic thickness-pitch metaphors causes Farsi speakers to use mental representations of spatial thickness to think about musical pitch, repeatedly using similar linguistic metaphors during training should cause Dutch speakers to perform similarly to Farsi speakers on the nonlinguistic thickness-interference task.

Method

Participants and materials. 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. Materials used were the same as in the thickness-interference task of Experiment 1, with the addition of 196 fill-in-the-blank sentences.

Procedure. During the training phase, participants completed 196 fill-in-the-blank sentences, presented on a computer screen. Participants in the thickness-training condition completed the sentences by typing the words *dunner* ("thinner") or *dikker* ("thicker"), and those in the height-training condition used the words *hoger* ("higher") or *lager* ("lower"). In both conditions, 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" or "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" or "A flute sounds thicker/thinner than a tuba"). Participants were left to infer the correct mapping based on three correctly completed example sentences and on the feedback they received after each trial, either *goed* ("correct") or *fout* ("incorrect"). Training took about 20 min. After the 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 task (M = 99%, SD = 1.05) and the thickness-training task (M = 99%, SD = 0.77). Accuracy did not differ between the tasks (difference = 0.2%), $t(58) = 0.96$, n.s.

Test phase. The thickness of the lines in the thickness-interference task influenced pitch reproduction in participants who underwent thickness training (mean slope = –1.45), $p = .003$, but not in those who had height training (mean slope = –0.08), n.s. The slope of the effect of normalized line thickness differed between the two conditions (difference = 1.38), $t(58) = 1.84$, $p = .04$, one-tailed (see also Fig. S3 in the Supplemental Material). The effect of line thickness on pitch reproduction in thickness-trained participants was statistically indistinguishable from the effect in native Farsi speakers in Experiment 1 (difference of slopes = 1.39), $t(48) = 1.12$, n.s., and was greater than the effect in untrained Dutch speakers in Experiment 1 (difference of slopes = 2.06), $t(48) = 2.02$, $p = .05$.

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

Experiment 4: Does Language Create Space-Pitch Mappings?

What role does language play in shaping nonlinguistic pitch representations? According to one proposal, language may be instrumental in creating mental metaphors (Boroditsky, 2001; Gentner & Wolff, 2000). Using spatial words like *high* or *thick* metaphorically in referring to pitch might encourage speakers to create analogical correspondences between space and pitch that did not exist (or were not used) prior to their exposure to the linguistic metaphors.

Alternatively, rather than creating new representations, linguistic metaphors might simply modulate the strength of preexisting mental metaphors (Casasanto, 2008). Studies have shown that prelinguistic infants are sensitive to the association between height and pitch found in English and Dutch metaphors (Walker et al., 2010), as well as to the association between thickness and pitch found in Zapotec and Farsi metaphors (Dolscheid, Hunnius, Casasanto, & Majid, 2012). Language, then, may serve to adjust the relative strengths of prelinguistic 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-metaphor language like Dutch would strengthen their height-pitch mapping at the expense of their thickness-pitch mapping, as a result of competitive correlational learning (Casasanto, 2008, 2010). The opposite would be true for speakers of a thickness-metaphor language like Farsi, whose linguistic experience would strengthen their preexisting thickness-pitch mapping at the expense of their preexisting height-pitch mapping.

To distinguish between these alternatives, in Experiment 4 we trained Dutch speakers to use a *reversed-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, training Dutch speakers to use the reversed-thickness mapping should be just as effective as training them to use the Farsi-like mapping because the two mappings are equally novel and systematic. Alternatively, if learning a new linguistic metaphor influences pitch representations by strengthening a preexisting space-pitch mapping, training should be more effective when the new metaphor corresponds to one of the space-pitch mappings found in prelinguistic infants than when it contradicts one of these mappings.

Method

Native Dutch speakers (N = 30) participated for payment. The materials and procedure were identical to those used in Experiment 3, with one exception: Participants were trained to use the reversed-thickness (reversed-Farsi) mapping (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 (M = 95%, SD = 6.69). Comparing accuracy between participants who received reversed-thickness training (Experiment 4) and those who received thickness training (Experiment 3), 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. The thickness of the lines in the thickness-interference task did not influence participants' pitch reproduction after reversed-thickness training (mean slope = –0.34), n.s. When we compared effects between participants who received reversed-thickness training (Experiment 4) and those who received thickness training (Experiment 3), we found that the effect of thickness on pitch reproduction was significantly greater in those who received thickness training (difference of slopes = 1.12), $t(58) = 2.20$, p = .03 (see also Fig. S4 in the Supplemental Material).

Results of the reversed-thickness training suggest that language use did not create a Farsi-like space-pitch mapping in Experiment 3. Rather, using Farsi-like linguistic metaphors strengthened the preexisting "low = thick, high = thin" mapping that was not evident in adult Dutch speakers' language or thought (see Experiment 1), but has been observed in prelinguistic infants (Dolscheid et al., 2012).

General Discussion

Dutch and Farsi speakers, who use different linguistic metaphors for pitch, form correspondingly different nonlinguistic pitch representations. We showed this through a double dissociation between the performance of Dutch and Farsi speakers on a pair of nonlinguistic psychophysical tasks. Dutch speakers, who talk about pitches as high (*hoog*) and low (*laag*), incorporated irrelevant height information into their pitch estimates (but ignored irrelevant thickness information). Farsi speakers, who talk about pitches as thin (*nāzok*) and thick (*koloft*), 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 that native Farsi speakers showed. In addition to demonstrating a language-thought correlation, our 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, which suggests that the 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 as a result of some form of competitive 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 received 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 testing, 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"

According to 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 on-line 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 perceptuomotor 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 on-line, overtly or covertly. The observed effect of language on thought even under verbal interference contrasts with other Whorfian effects reported previously—for example, effects in the domain of color (e.g., Winawer et al., 2007)—which disappear under verbal interference.

Origins of space-pitch mappings

Prelinguistic infants appear to be sensitive to both height-pitch and thickness-pitch mappings (Dolscheid et al., 2012; Walker et al., 2010), which suggests 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). Our results are consistent with this proposal, as training adults to use new verbal metaphors was more effective in changing their nonlinguistic pitch representations when the new metaphors corresponded to preverbal 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 reversed-Farsi 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? One possibility is that both the height-pitch and the thickness-pitch mappings reflect innate cross-modal correspondences that have no experiential basis (Walker et al., 2010). Alternatively, both of these space-pitch mappings could 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 the mind emerge over developmental time, as individuals track experiential regularities (Lakoff & Johnson, 1999), and to what extent these regularities emerged over evolutionary time, as the neural substrates of spatial cognition were exapted for nonspatial functions (Pinker, 1997).

Conclusion

Whatever their origins may be, different space-pitch mappings have become encoded in the languages people speak, in expressions so highly conventionalized that 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.

Acknowledgments

The authors thank Ozge Ozturk, Francisco Torreira, and Tomas Bergvelt.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

This research was funded by the Max Planck Gesellschaft, by Grant P09-SEJ-4772 from the Junta de Andalucía/European Regional Development Fund, and by a James S. McDonnell Foundation Scholar Award to D. C.

Supplemental Material

Additional supporting information may be found at http://pss.sagepub.com/content/by/supplemental-data

Notes

1. All Farsi-speaking participants spoke Farsi on a daily basis, according to their responses to a questionnaire. Most also spoke some Dutch or English and so may have been exposed to linguistic height metaphors. In principle, such exposure could eliminate the predicted effects, but it could not produce them.
2. Although it is generally assumed that language is not playing an on-line 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).

References

Ashley, R. (2004). Musical pitch space across modalities: Spatial and other mappings through language and culture. In P. W. S. Lipscomb, R. Ashley, & R. Gjerdingen (Eds.), *Proceedings of the 8th International Conference on Music Perception and Cognition* (pp. 64–71). Adelaide, Australia: Causal Productions.

Boroditsky, L. (2001). Does language shape thought? Mandarin and English speakers' conceptions of time. *Cognitive Psychology*, *43*, 1–22.

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

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

Casasanto, D., & Bottini, R. (2010). Can mirror-reading reverse the flow of time? In N. S. Hölscher, T. F. Shipley, M. Olivetti Belardinelli, & J. A. Bateman (Eds.), *Spatial cognition VI* (pp. 335–345). Berlin, Germany: Springer.

Casasanto, D., & Chrysikou, E. G. (2011). When left is "right": Motor fluency shapes abstract concepts. *Psychological Science*, *22*, 419–422.

Dolscheid, S., Hunnius, S., Casasanto, D., & Majid, A. (2012). The sound of thickness: Prelinguistic infants' associations of space and pitch. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th annual meeting of the Cognitive Science Society* (pp. 306–311). Austin, TX: Cognitive Science Society.

Douglas, K. M., & Bilkey, D. K. (2007). Amusia is associated with deficits in spatial processing. *Nature Neuroscience*, *10*, 915–921.

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

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*, 333–336.

Gennari, S. P., Sloman, S., Malt, B. C., & Fitch, W. T. (2002). Motion events in language and cognition. *Cognition*, *83*, 49–79.

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: Erlbaum.

Lakoff, G., & Johnson, M. (1980). The metaphorical structure of the human conceptual system. *Cognitive Science*, *4*, 195–208.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to Western thought*. New York, NY: Basic Books.

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*, 1189–1207.

Lupyan, G. (2012). Linguistically modulated perception and cognition: The label-feedback hypothesis. *Frontiers in Cognition*, *3*, 54. Retrieved from http://www.frontiersin.org/Cognition/10.3389/fpsyg.2012.00054/full

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

Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition*, *108*, 155–184.

Pinker, S. (1997). *How the mind works*. New York, NY: Norton.

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

Shayan, S., Ozturk, O., Bowerman, M., & Majid, A. (2012). *Spatial metaphor in language can promote the development* *of cross-modal mappings in children*. Manuscript submitted for publication.

Shayan, S., Ozturk, O., & Sicoli, M. (2011). The thickness of pitch: Cross-modal metaphors in Farsi, Turkish and Zapotec. *The Senses & Society*, *6*, 19–29.

Slobin, D. I. (1996). From “thought and language” to “thinking for speaking.” In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70–96). Cambridge, England: Cambridge University Press.

Stone, R. M. (1981). Toward a Kpelle conceptualization of music performance. *Journal of African Folklore*, *94*, 188–206.

Tillmann, B., Jolicoeur, P., Ishihara, M., Gosselin, N., Bertrand, O., Rossetti, Y., & Peretz, I. (2010). The amusic brain: Lost in music, but not in space. *PLoS ONE*, *5*, e10173. Retrieved from http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0010173

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

Walker, P., Bremner, J. G., Mason, U., Spring, J., Mattock, K., Slater, A., & Johnson, P. (2010). Preverbal infants’ sensitivity to synaesthetic cross-modality correspondences. *Psychological Science*, *21*, 21–25.

Winawer, J., Witthoft, N., Frank, M., Wu, L., Wade, A., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences, USA*, *19*, 7780–7785.