Pitch Affects Estimates of Space but not *Vice Versa*

Alexander Kranjec (kranjeca@duq.edu)
Psychology Department, Duquesne University
Center for the Neural Basis of Cognition, Carnegie Mellon University
Pittsburgh, PA, USA

Matthew Lehet (mil@andrew.cmu.edu)
Psychology Department, Carnegie Mellon University
Pittsburgh, PA, USA

Anjan Chatterjee (anjan@mail.med.upenn.edu)
Neurology Department and Center for Cognitive Neuroscience, University of Pennsylvania
Philadelphia, PA, USA

Abstract

The idea that we think about relatively abstract domains (like time) in terms of more concrete domains (like space) but not *vice versa* can be traced to conceptual metaphor theory. Experiments using verbal and/or visual stimuli suggest a deep ontological basis for space-time asymmetries. Yet vision makes a privileged contribution to spatial processing raising questions about modality. Recently, we found that *in sound*, time and space are mutually contagious, with a larger effect of time on space. Here we examine the mutual effects of space, time, and *pitch*, a uniquely auditory attribute. If space is more abstract than time in sound, space should be *more* easily contaminated by pitch, while being *less* effective in contaminating it. While time and pitch were shown to be mutually contagious, pitch affected estimates of space but not *vice versa*. Results overall suggest that in sound, time is not fundamentally more abstract than space.

Keywords: space and time; language and thought; metaphor; embodiment

Introduction

Time is talked about in spatial terms much more frequently than space is talked about in terms of time. In many ways time *must* be talked about using the language of space, whereas the opposite is not true. Such asymmetrical linguistic patterns have been interpreted to suggest an asymmetry at a deeper level of conceptual organization. According to *conceptual metaphor theory* (Lakoff & Johnson, 1999) we think about relatively abstract domains (like time) in terms of more concrete domains (like space). It is maintained that this organizational principle serves the functional role of making more abstract concepts easier to talk about *and think* about. It is argued that this is necessary because we can directly see and touch things "in space" in a way that we cannot "in time". This suggests that thinking about time in terms of space runs cognitively deep, and reflects a mental organization more fundamental than that observed at the relatively superficial level of words.

Casasanto and Boroditsky (2008) originally provided experimental evidence for this theoretical organizational principle. They were interested in whether the asymmetry of space-time metaphors in language predicted a similar asymmetry in visual perception. Specifically, they found that the remembered size of a line in space concordantly modulated recall for its duration, but not *vice versa*. That is, (spatially) longer lines were remembered as being presented for longer times, but lines of greater durations were not remembered as having greater spatial extent. The results were consistent with the idea that asymmetrical patterns of space-time mappings in language are preserved further down at the level of perception. However, because this study relied on visual tasks, it was still not clear if observed behavioral asymmetries between time and space reflect (1) general ontological (or even metaphysical) relations dependent on each domain's relative level of "abstractness" or (2) a less general, modality-specific contribution of visual representations in humans.

There are intuitive reasons to think that time−space asymmetries observed in vision might actually be reversed in sound, as time, more than space, seems to be an intimate part of our auditory experience. For example, whereas spatial relations and visual objects tend to be persistent, sound, like time, is relatively transient (Galton, 2011). Temporal information is more critical and/or salient in common forms of experience grounded in sound perception (e.g., music and speech) and sound localization

is less precise than object localization in vision (Kubovy, 1988). And neurophysiologically, while the retina preserves analog spatial relations in early representations, the cochlea does not (Moore, 1977; Ratliff & Hartline, 1974). Thus, one might argue that relations between sound and time are relatively more concrete than relations between sound and space.

To distinguish between domain-general and domain-specific explanations for prior experimental results, a recent study directly probed time–space relations in the auditory domain. Kranjec, Lehet, and Chatterjee (2013) employed a task that closely followed Casasanto and Boroditsky (2008) but used auditory instead of visual stimuli finding that space and time are mutually contagious. Furthermore, as predicted by the privileged relation between auditory and temporal processing, the perceived duration of a stimulus had a larger effect on perceived spatial displacement than the reverse. This asymmetry runs in the opposite direction of the asymmetry found in the visual domain as predicted by patterns of language use (Casasanto & Boroditsky, 2008) and the relatively spatial nature of vision.

While these prior results are suggestive of a perceptual asymmetry running opposite to that observed in the visual domain, broader claims regarding a deep ontological asymmetry between time and space in the auditory domain are still currently unwarranted. Although "in sound," time appeared to influence judgments of spatial displacement more than *vice versa*, these results may not generalize. To further probe the relative abstractness of space and time in sound, the present study examines the mutual effects of space, time, and *pitch*, a uniquely auditory attribute. It was reasoned that if space is more abstract than time in sound, space should be *more* easily contaminated by pitch, while being *less* effective in contaminating it.

Space, Time, and Pitch

The perception of pitch makes possible the processing of melody in music, and prosody in speech. Defined as the perceived frequency or "repetition rate of an acoustic waveform" (Oxenham, 2012) pitch is, together with loudness and timbre, one of three basic auditory sensations. Current theories suggest that properties of the physical stimulus and the physiological mechanisms for transduction and neural representation, in addition to prior experience, all play a significant role in pitch perception. This most likely involves both temporal and place coding throughout the auditory system. When sound enters the cochlea, the distinct frequencies that make up an acoustic waveform activate tuned neural sites arranged along its membrane in an analog manner. Such *tonotopic*, "rate-place" (or time-space) mapping is preserved in the auditory processing system as far as the primary auditory cortex. (See Oxenham (2012) for a review.) As such, pitch perception involves the representation of both spatial and temporal information at multiple levels of processing. The centrality and salience of pitch perception in auditory experience, and its fundamental spatiotemporality make it an ideal domain for further testing hypotheses supported by our previously reported (Kranjec, et al., 2013) research.

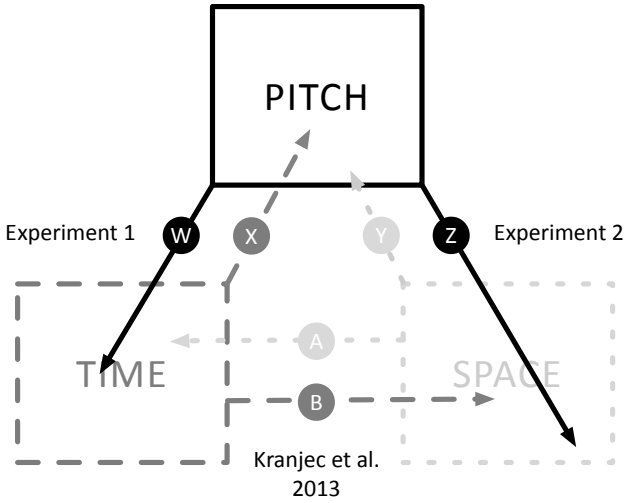

Figure 1. Space, Time, and Pitch in Sound. The solidity and darkness of each box indicates that domain's relative level of concreteness in the auditory modality. (I.e., in sound PITCH is more concrete than TIME, which is more concrete than SPACE.) The extent that an arrow extending from a particular domain penetrates another domain suggests how one domain is expected to influence another. (I.e., more concrete domains are expected to exert more influence on relatively abstract domains, while being less affected by them.) Kranjec et al. (2013) found that time and space are mutually contagious, with a larger effect of time on space (B>A). The current study investigates relations between space, time, and pitch. As suggested by the diagram, in Experiments 1 & 2 PITCH is expected to exert a *greater influence* on SPACE as compared to TIME (Z>W or *PITCH*→ SPACE>TIME), but be *less affected* by SPACE as compared TIME (Y<X or SPACE<TIME →*PITCH*). In general, PITCH is expected to exert more influence on SPACE and TIME than *vice versa* (W+Z>X+Y or *PITCH*→ > →*PITCH*).

The current experiments test the extent that irrelevant spatial and temporal information contaminate representations of pitch (and vice versa). Here we observe the amount of cross-domain interference between time and pitch (Exp. 1) and space and pitch (Exp. 2) using the same design, analyses, and logic of our prior (2013) study (and Casasanto & Boroditsky, 2008). If space is more abstract than time in sound then, on the one hand, pitch should affect spatial judgments *more* than temporal judgments (*PITCH*→ **SPACE>**TIME), but on the other hand, space should be *less* effective than time in influencing pitch judgments (**SPACE<**TIME →*PITCH*).

Experiments

General Procedure

The general design of both Experiments 1 and 2 was identical to that of a previously reported experiment (Kranjec, et al., 2013). Participants were equipped with headphones and seated at a computer for a self-paced experiment. Participants initiated the beginning of each new trial and the start of each within-trial component. Each trial consisted of two sounds, a *target sound* followed by a *playback sound*. In the first part of each trial, the target sound was presented, and participants were instructed to attend to *either* the duration and pitch of the stimulus (Experiment 1) *or* the distance and pitch of the stimulus (Experiment 2). After attending to the target sound participants were informed of the trial type and instructed to press the spacebar to begin the playback sound. The playback sound provided the medium for the participant to reproduce either the duration, displacement or pitch depending on the experiment and trial type. All stimuli were created using Matlab and played using the OpenAL library provided with Psychophysics Toolbox extensions (Brainard, 1997).

Participants

Forty-two members of the University of Pennsylvania community participated for payment. All participants were right-handed, native English speakers, and between 18-26 years of age. Twenty-two participants performed Experiment 1. Data from two of these participants were excluded from the final analyses because their reaction times across conditions were greater than two standard deviations from the mean. Twenty distinct participants performed Experiment 2.

Experiment 1: Time and Pitch

In Experiment 1, when the target sound was presented, participants were instructed to attend to both the *duration and pitch* of the stimulus. The target sound in Experiment 1 was a sound consisting of a variable and continuous range of frequencies presented over a variable period of time in both ears. Target sounds were of nine durations [lasting between 1000 and 5000ms with 500ms increments as in Kranjec, et al. (2013)] and 9 frequencies (ranging between 150 and 1350Hz in increments of 150Hz). All durations and frequencies were crossed to create 81 distinct target sounds. Each discrete stimulus was used twice, once in the duration condition and once in the pitch condition. The initial frequency of the target sound began within the higher (2250Hz) or lower (990Hz) ends of the audible range of speech with a randomized jitter between 1 and 50Hz. Frequency endpoints were determined by varying the number of frequency increments the sound moved through across trials. Frequency "direction" (high to low, or low to high) was random across trials.

After attending to the target sound, participants in Experiment 1 were informed of the trial type (duration or pitch) and instructed to press the spacebar to begin the playback sound. The playback sound provided the medium for the participant's response. It presented the same frequency ranges in the opposite direction, starting at the frequency endpoint of the target sound and moving towards the start point and lasted for a maximum of 8.5 seconds or until the participant ended the trial by responding. On *duration* trials, participants were instructed to respond when the playback sound duration was equal to the target sound duration. On *pitch* trials, participants were instructed to respond when the playback sound span equaled that of the target sound's frequency range. For all trials, there were at least 5 additional frequency increments and 7 additional duration increments within the playback sound to allow participants the possibility to both overshoot and undershoot their estimates. Data for both duration and frequency judgments were collected regardless of condition.

Experiment 2: Space and Pitch

The procedure for Experiment 2 was identical to that in 1 but with *distance* replacing duration as a domain of interest. In Experiment 2 participants were instructed to attend to both the *distance and pitch* of the stimulus. Target sounds were of nine distances [moving between .5 and 4.5m in increments of .5m, as in (Kranjec, Lehet, et al., 2013)] and 9 frequencies (ranging between 150 and 1350Hz in increments of 150Hz as in Experiment 2A) all crossed to create 81 discrete stimuli. The initial location of the target sound was an average of 2.75m to the left or right of the listener with a jitter of between .1 and .5m. Starting locations on the right indicated leftward moving trials and starting locations on the left indicated rightward moving trials. Starting locations were randomly assigned to stimuli with an even number of right and leftward moving trials. The plane of movement was 1 meter in front of the listener. Stimuli were created using Matlab and played using the OpenAL library provided with Psychophysics Toolbox extensions (Brainard, 1997). The OpenAL library is designed to model sounds moving in virtual metric space for a listener wearing headphones.

In Experiment 2, the playback sound began in the final spatial location and frequency endpoint of the preceding target sound and moved in the reverse direction (both in terms of space and pitch). Directionality in space (left to right or right to left) and pitch (high to low or low to high) was randomized across all trials. On *distance* trials, participants were instructed to respond when the playback

sound reached the start location of the target sound. In this manner, the participant's head provided a fixed reference point for judging distance. On *pitch* trials, participants were instructed to respond when the playback sound spanned the target sound's frequency range.

Results: Experiments 1 and 2

Between Experiments 1 and 2 there are 4 main correlations to consider. They describe the effects of frequency on (A) distance estimates (PITCH→ SPACE) and (B) duration estimates (PITCH→ TIME) and the effects of (C) distance and (D) duration on frequency estimates (SPACE →PITCH and TIME →PITCH, respectively). These results are displayed in Figure 2. A comparison of r^2 values between conditions/experiments is depicted in Figure 3.

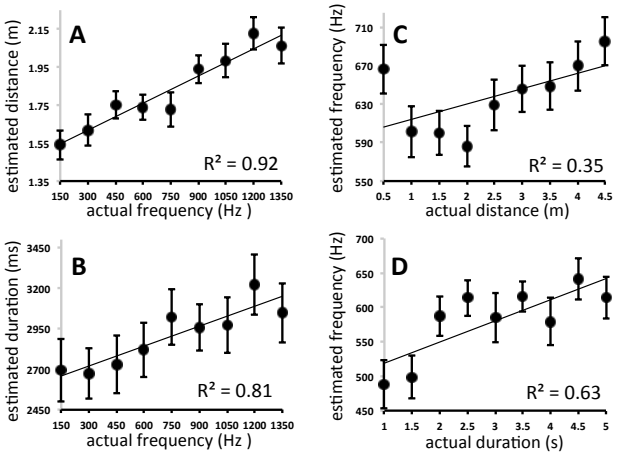

Figure 2. Results for Experiments 1 and 2. Because all 9 intervals used for each domain were fully crossed in Experiments 1 and 2, the expected average for estimates across all participants for a particular trial type (distance, duration, or frequency estimation; y-axis) can be described as the average of all 9 interval values for that domain presented at each interval of the irrelevant distractor domain (actual frequency, distance, or duration; x-axis). If the irrelevant domain on *x* exerted no influence on estimation for *y* one would expect a horizontal line. Deviation from that horizontal represents cross-domain interference. (A) Effect of frequency on distance estimates (expected= 2.5m at each interval of actual frequency). (B) Effect of frequency on duration estimates (expected= 3000ms at each interval of actual frequency). (C) Effect of distance on frequency estimates (expected= 750Hz at each interval of actual distance). (D) Effect of duration on frequency estimates (expected= 750Hz at each interval of actual duration). Error bars refer to standard error of the mean.

Analyses demonstrate that actual frequency (PITCH→) affected estimates of spatial displacement (Fig. 2A: y = 0.0005x + 1.4745, r^2 = .92, df = 7, p < .001) and duration (Fig. 2B: y = 0.4098x + 2597.1, r^2 = .81, df = 7, p = .001). The effect of distance on frequency estimation was not significant (Fig 2C: y=15.955x + 598.21, r^2= .35, df = 7, p = .09), while actual duration affected estimates of frequency (→PITCH) (Fig. 2D: y = 30.7x + 488.22, r^2= .63, df = 7, p = .01). The effect of actual frequency on spatial displacement (r^2 = .92, Fig. 2A) was significantly greater than the effect of space on frequency estimation (r^2= .35, Fig. 2C). (Difference of correlations = 0.57; z = -2.12 one-tailed, p = .01). Correlation coefficients for PITCH→ TIME (r^2 = .81, Fig. 2B) and TIME →PITCH (r^2= .63, Fig. 2D) effects were not significantly different from one another.

Participants' overall estimates of duration, spatial displacement, and pitch were accurate. The effects of actual duration on estimated duration (y = 187.04x + 2122 r² = 0.88, df = 7, p < .001), actual frequency on estimated pitch (*Exp. 1*: y = 0.2555x + 431.53 r² = 0.91, df = 7, p < .001), actual spatial displacement on estimated displacement (y = 0.4874x + 0.6134 r² = 0.98, df = 7, p < .001), and actual frequency on estimated pitch (*Exp. 2*: 0.4425x + 306.19 r² = 0.99, df = 7, p < .001) were all highly reliable but not significantly different from one another.

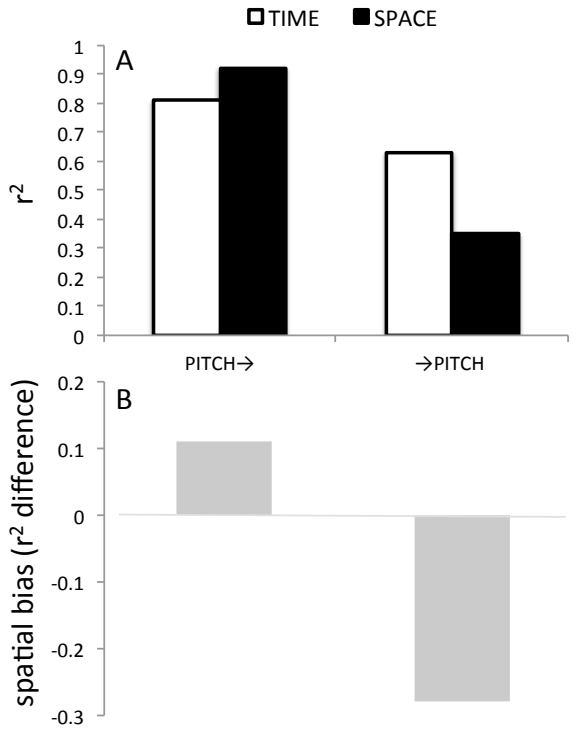

Figure 3. Comparison of r^2 values and difference scores for Experiments 1 and 2. (A) r^2 values for the effects of pitch (PITCH→) on time and space and the effects of time and space on pitch (→PITCH). (B) "Spatial bias" is the r^2 difference score (space – time) for both types of pitch trials (PITCH→ and →PITCH) illustrating the relative extent that space is *modulated by*, or is effective *in modulating*, pitch as compared to time.

Discussion

We predicted that if space is more abstract than time in sound then pitch should affect spatial judgments *more* than temporal judgments (PITCH→ **SPACE>**TIME: a positive spatial bias), but that space should be *less* effective than time in influencing pitch judgments (**SPACE<**TIME →PITCH: a negative spatial bias). The significant asymmetry in the effects of pitch-on-space vs. space-on-pitch, together with an inspection of the r^2 difference scores (Fig. 3B) is consistent with this prediction. As compared to time, there is a positive spatial bias when pitch is serving as a modulator (PITCH→ SPACE), and a negative spatial bias when space is provided the opportunity to affect pitch (SPACE →PITCH). The pattern of results suggests that in sound, space is particularly sensitive to irrelevant information while being less effective in modulating other kinds of information. This is the profile one would expect from a more abstract representation with a relatively fragile cognitive organization.

The asymmetry reported here is also predicted by the temporal nature of auditory processing. This asymmetry runs in the opposite direction to that found in the visual domain as predicted by patterns of language use (Casasanto & Boroditsky, 2008) and the relatively spatial nature of vision. More generally, the results do not support the idea that *time is more abstract than space* at the level of general ontology and/or basic cognitive architecture. Rather, they suggest that relations between space and time may be more or less abstract depending on the sensory modality in which particular stimuli are processed or experienced.

The results are also consistent with a prior study that used an analogous design to investigate space-time relations directly. Kranjec et al. (2013) found space and time to be mutually contagious with the perceived duration of a stimulus having a larger effect on perceived displacement than vice versa. Taken together, the results from both studies support a view of embodied cognition that takes into account the contributions of a particular sensory modality in processing the abstract qualities of a stimulus. While space and time can both be considered relatively abstract concepts, relations between objects as experienced in either (whether seen or heard) may be more or less so depending on a range of species-specific and contextual variables.

The general idea that visuospatial representations are central to how people talk and think is well established (Chatterjee, 2001; Johnson-Laird, 1986; Talmy, 2000; Tversky, 2005). For humans, "embodied spatial representations" important for structuring other forms of thought and language are likely visuospatial in nature. Because humans have a general visual bias in perception, communication, and neural organization, there may be a tendency for us to experience space as relatively less abstract than time. But this does not mean that space is necessarily less abstract than time, or that other organisms experience space and time as we do. While it is famously difficult to imagine the quality of conscious experience in another organism (Nagel, 1974) perhaps it is the case that animals (like bats) which rely more on audition than vision to find objects in a dynamic environment could be biased to experience time as less abstract than space.

A more tractable issue worth reconsidering concerns the question of why time is generally assumed to be more abstract than space in the first place. The argument may be based on the idea that time, as compared to space, cannot be "directly perceived" (Ornstein, 1969), or that we cannot "see or touch" time (Casasanto, Fotakopoulou, & Boroditsky, 2010). Yet there are known, widely distributed, neural mechanisms specific to temporal processing, and little basis for the assumption that spatial relations are themselves perceived directly (Kranjec & Chatterjee, 2010). The experience of space and time both involve inherently relational processes, making the representation of both relatively abstract.

For example, processing locations between objects in an array using vision is arguably no more or less direct than processing rhythm in a sequence of beats using audition, with each requiring the representation of a number of abstract *relations* between objects or sounds. That is, there is no reason to think that we can directly "see" space any more than we can "hear" time. Nowhere is the dissociation between vision and spatial processing more apparent than in *simultanagnosia*, a neuropsychological condition in which patients are characteristically unable to perceive more than a single object despite having intact visual processing (Kranjec, Ianni, & Chatterjee, 2013; Luria, 1959). Nonetheless, visuo-spatial and audio-temporal relations appear to be privileged. Privileged relations between particular sensory modalities and experiential domains may play some part in determining what we come to label abstract or concrete. Further research is needed to determine why some senses are subjectively felt to be more or less abstract than others, and the specific roles that spatial and temporal organization play in structuring our sensory experience.

References

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial vision, 10*(4), 433-436.

Casasanto, D., & Boroditsky, L. (2008). Time in the mind: Using space to think about time. *Cognition, 106*(2), 579-593.

Casasanto, D., Fotakopoulou, O., & Boroditsky, L. (2010). Space and Time in the Child's Mind:

Evidence for a Cross-Dimensional Asymmetry. *Cognitive Science, 34*(3), 387-405.

Chatterjee, A. (2001). Language and space: some interactions. *Trends in Cognitive Science, 5*, 55-61.

Galton, A. (2011). Time flies but space does not: Limits to the spatialisation of time. *Journal of Pragmatics, 43*(3), 695-703.

Johnson-Laird, P. N. (1986). *Mental models: Towards a cognitive science of language, inference, and consciousness*: Harvard Univ Pr.

Kranjec, A., & Chatterjee, A. (2010). Are temporal concepts embodied? A challenge for cognitive neuroscience. *Frontiers in Psychology, 1*(240), doi: 10.3389/fpsyg.2010.00240.

Kranjec, A., Ianni, G., & Chatterjee, A. (2013). Schemas reveal spatial relations to a patient with simultanagnosia. *Cortex, 49*(6), 1983-1988.

Kranjec, A., Lehet, M., & Chatterjee, A. (2013). *Space and Time are Mutually Contagious in Sound.* Paper presented at the 35th Annual Conference of the Cognitive Science Society.

Kubovy, M. (1988). Should we resist the seductiveness of the space: time:: vision: audition analogy?

Lakoff, G., & Johnson, M. (1999). *Philosophy in the Flesh*. New York, NY: Basic Books.

Luria, A. (1959). Disoders of Simultaneous Perception in a Case of Bilateral Occipto-Parietal Brain Injury. *Brain, 82*(3), 437-449.

Moore, B. C. (1977). *Introduction to the psychology of hearing*. Baltimore: University Park Press.

Nagel, T. (1974). What is it like to be a bat? *The philosophical review, 83*(4), 435-450.

Ornstein, R. E. (1969). *On the experience of time*: Penguin.

Oxenham, A. J. (2012). Pitch perception. *The Journal of Neuroscience, 32*(39), 13335-13338.

Ratliff, F., & Hartline, H. K. (1974). *Studies on Excitation and Inhibition in the Retina: A collection of Papers from the Laboratories of H. Keffer Hartline*: Rockefeller Univ. Press.

Talmy, L. (2000). *Towards a cognitive semantics: Concept structuring systems*. Cambridge, MA: The MIT Press.

Tversky, B. (2005). Visuospatial reasoning. *The Cambridge handbook of thinking and reasoning*, 209-249.