The Sound of Time: Cross-Modal Convergence in the Spatial Structuring of Time.

Daniël Lakens¹
Gün R. Semin¹
Margarida V. Garrido²

In Press, Consciousness and Cognition

¹ Faculty of Social and Behavioral Sciences, University of Utrecht, Netherlands
² ISCTE, Lisbon University Institute

Word count: 3306 words

KEYWORDS: Abstract Concepts; Time; Grounded Cognition; Auditory Modality; Representation
Abstract

In a new integration, we show that the visual-spatial structuring of time converges with auditory-spatial left-right judgments for time-related words. In Experiment 1, participants placed past and future-related words respectively to the left and right of the midpoint on a horizontal line, reproducing earlier findings. In Experiment 2, neutral and time related words were presented over headphones. Participants were asked to indicate whether words were louder on the left or right channel. On critical experimental trials, words were presented equally loud binaurally. As predicted, participants judged future words to be louder on the right channel more often than past related words. Furthermore, there was a significant cross-modal overlap between the visual-spatial ordering (Experiment 1) and the auditory judgments (Experiment 2), which were continuously related. These findings provide support for the assumption that space and time have certain invariant properties that share a common structure across modalities.

Word count: 146 words
The Sound of Time: Cross-Modal Convergence in the Spatial Structuring of Time.

How do we represent and think about abstract concepts that do not afford sensorimotor experiences? This question has been at the heart of recent discussions in embodied approaches to cognition (cf. Barsalou, 2008; Dove, 2009; Glenberg et al., 2008; Mahon & Caramazza, 2008). One answer, based on conceptual metaphor theory (Lakoff & Johnson, 1999), proposes that thoughts about abstract concepts, such as time, are structured by perceptual experiences, such as space (Boroditsky, 2000; Tversky, Kugelmass, & Winter, 1991). Indeed, many studies have shown that spatial information can influence time-related judgments (Boroditsky, 2000, 2001; Boroditsky & Ramscar, 2002; Casasanto & Boroditsky, 2008) or the categorization of time-related words (Ouellet, Santiago, Funes, & Lupiáñez, 2010; Santiago, Lupiáñez, Pérez, & Funes, 2007; Torralbo, Santiago, & Lupiáñez, 2006).

Recent work has revealed the intricate subtleties through which the cognitive representation of time is inherently intertwined with the representation of space. Bimanual response tasks have revealed compatibility effects between time-related stimuli and the spatial position of response keys. For example, when participants were asked to indicate whether the last cue in a sequence of eight auditory cues (of which the first seven were played with a 500 ms interval) was presented earlier or later than 500 ms, responses were faster when the left (vs. right) key was used to respond to auditory cues presented earlier than 500 ms, and the right (vs. left) key was used to respond to auditory cues presented later than 500 ms (Ishihara, Keller, Rossetti, & Prinz, 2008). Similar stimulus-response compatibility effects have also been observed in other studies. When participants were asked to indicate the duration of a cross presented on the screen, short
(vs. long) durations were classified faster with the left (vs. right) response key, relative to the reversed stimulus-response mapping (Vallesi, Binns, & Shallice, 2008; Vallesi, McIntosh, & Stuss, in press). Comparable stimulus-response compatibility effects have been reported for stimuli referring to actors born in the first or second half of the 20\textsuperscript{th} century that had to be categorized on date of birth by pressing the left or right response key (Weger & Pratt, 2008), or when past and future words presented auditorily to the left or right ear had to be categorized on temporal meaning (Ouellet, Santiago, Isreali, & Gabay, in press).

In addition to stimulus-response compatibility effects of the temporal meaning of words on left or right key presses, some interesting studies revealed cross-modal response interference effects in the domain of time. Dormal, Seron, and Pesenti (2006) observed that temporal judgments of the duration of a series of dots presented on the screen were facilitated when amount and duration of the stimuli were congruent, but interfered with duration judgments when the stimuli were incongruent. Building on these findings, research by Xuan, Zhang, He, and Chen (2007) revealed that error rates of temporal judgments for stimuli presented for a range of durations in a Stroop-like interference paradigm were affected by the brightness or size of the stimuli that was manipulated orthogonally, such that larger or brighter stimuli were judged to be presented for a longer duration. These studies provide support for the assumption that temporal information and non-temporal ordinal dimensions are related. The common representation of time and space is argued to fulfill the need for analogue sensorimotor transformations between space, time and other dimensions in order to coordinate actions (Walsh, 2003). Recent neurological evidence indicates that the right parietal cortex plays an important role in
temporal judgments of both visual and auditory durations (Bueti, Bahrami, & Walsh, 2008), and an integrative review supports the idea of domain-general representations of time, space, numbers, and other dimensions (Cohen Kadosh, Lammertyn, & Izard, 2008; see also Cohen Kadosh & Walsh, 2009).

The issue of whether the cross-modal interactions between temporal and spatial stimuli are due to the representation of time flowing from left to the right, or a facilitation of response codes has been addressed in recent research (see Ouellet, Santiago, Funes, & Lupiáñez, 2010; Vallesi et al., in press). This question has been investigated using a variety of response time paradigms. Some studies have led to the conclusion that the observed effects were primarily (though not exclusively) due to response facilitation, for example by comparing a bimanual categorization task (where participants have to indicate whether a stimulus appears on the left or right side of the screen by pressing the left or right response key) with a stimulus detection task (responding with a single key press when a stimulus appears on the screen (Weger & Pratt, 2008, see also Ouellet, et al., in press). Recently however, studies have provided initial support for the assumption that future and past-related words focus people’s attention to the right and the left, independent of response congruency effects (Ouellet et al., 2010). These findings provide preliminary support for the idea that temporal information not only primes response codes, but also influences the orientation of spatial attention, in line with the idea that time is visually represented from left-to-right in space.

The current research explores the novel question whether the visual and auditory mappings of time in space converge, by investigating whether auditory judgments for time-related words showed a comparable spatial bias as visual judgments. We propose
that a multimodal representation of time in space should give rise to considerable overlap between how time is structured in visual space, and how time is structured in auditory space. If the future is represented more to the right in auditory space (as it is in visual space), binaurally presented future-related (vs. past-related) words should be judged to be louder on the right (vs. left) auditory channel. We investigated whether a graded left-to-right positioning of time-related words in visual space would significantly overlap with auditory-spatial judgments, in a paradigm which does not rely on response interference effects. By examining the linear relationship between the visual and auditory structuring of time, the current studies aim to provide support for a graded multi-modal representation of time in horizontal space. Moreover, these studies were also designed to examine whether the spatial structuring of past and future related words represents a continuum driven by temporal meaning, where the word ‘past’ is placed further to the left than ‘yesterday’, or whether the past and future are dichotomously structured in space, with past to the left and future to the right, without a grading of words such as ‘past’ and ‘yesterday’ in space.

The first experiment laid the groundwork for the investigation of the cross-modal convergence of the auditory and visual-spatial structuring of time, by asking participants to explicitly place words referring to the past and the future on a horizontal line. For the second experiment, we designed a novel auditory judgment task where participants were asked to indicate whether stimuli presented binaurally over headphones were louder on the right or left channel. On critical trials, words were presented equally loud on both channels. These critical trials were distributed within a number of randomly presented filler trials, with words varying systematically in their volume on the right and left
auditory channels.

In line with the existence of a mental time line where the past is on the left, and the future on the right, and based on the fact that the auditory modality is highly sensitive to spatial information (Blauert, 1997), we expected that time-related words which participants placed more on the right in a visual positioning task, would also be mapped further to the right in an auditory judgment task. Crucially, we predicted that participants would judge future words to be louder in the right ear more often than past-related words, whereby judgments on critical experimental trials constituted the chief dependent variable. Comparisons of the data from the first and second experiment permitted us to examine the cross-modal convergence in the structuring of time.

Experiment 1: Left-right visual positioning judgments

In this experiment, participants were asked to place past and future-related words on a horizontal line. We predicted that future-related words would be placed on the right half of the horizontal line, and past-related words on the left half of the line, replicating earlier findings in the literature (e.g., Tversky et al., 1991).

Participants. Fifty-six students (38 females, mean age 20) at Utrecht University participated in this study for payment or partial course credit.

Stimuli. The temporal stimuli consisted of eight past (past, day before yesterday, earlier, been, before, yesterday, recently, a moment ago) and eight future related words (later, future, day after tomorrow, coming, tomorrow, immediately, soon, shortly), which were selected so that they encompassed a continuous range of words related not only to the far past and future, but also referring to the immediate past and future. The temporal meaning of each word was further established by asking an independent sample of
eighteen participants to indicate the moment in time each stimulus word referred to on a 
9-point scale ranging from 1 (distant past) through 5 (present) to 9 (distant future). An 
analysis over items revealed that the temporal difference between past ($M = 2.83, SD = 
1.01$) and future ($M = 6.74, SD = 0.85$) referent words was highly significant, $t(14) = 
8.37, p < .001$. 

Procedure. We asked participants to place eight past and eight future-related words 
on a horizontal line, with an indicator at the midpoint of the line anchoring the present. 
The ends of the line were unmarked. The 16 time-related stimuli were presented in a 
random order. Participants’ task was to click on the line with a mouse to mark the 
position they thought best suited the stimulus. Responses were scored on a scale ranging 
from 0 (completely left) to 100 (completely right). 

Results and Discussion 

For the analyses we derived two averages of spatial position scores, one for future 
and one for past-related words. As expected, participants placed future-related words 
further to the right ($M = 66.42, SD = 3.99$) of the midpoint (scored as 50 on the 100-point 
scale) than past-related words ($M = 29.71, SD = 4.31$), $t(55) = 37.43, p < .001$. The 
average absolute distances from the midpoint of the horizontal scale (50) were calculated 
in an analysis over items for past ($M = 20.29, SD = 14.29$) and future ($M = 16.42, SD = 
12.69$) related words. These differences did not differ statistically, as revealed by an 
independent-samples $t$-test, $t(14) = .57, p = .58$, indicating that past and future words 
were equally distant from the scale midpoint. 

We conducted additional analyses to establish whether the spatial positioning of the 
stimuli is driven by their temporal meaning, and to test whether the spatial position of
past and future related words is dichotomous or continuous in nature. A linear regression analysis on the horizontal position of the time-related words in Experiment 1 with the temporal ratings provided by an independent sample as predictor revealed that, as expected, the differences in the temporal meaning of the stimuli predicted their horizontal spatial position, $\beta = .97$, $t = 14.56$, $p < .001$. Further analyses revealed no moderation of the dichotomous temporal category (past vs. future, dummy coded with past as the reference group), $\beta = .21$, $t = -1.39$, $p = .66$. This indicates that the horizontal position of the stimuli in Experiment 1 is continuously related to their temporal meaning, rather than being a simple dichotomous distinction between the categories of past and future.

In line with earlier findings (e.g., Tversky et al., 1991), the results show unambiguously that the positioning of past and future-related words show a left to right ordering on a horizontal dimension and suggest that the spatial structuring of past and future related words represents a continuum driven by their semantic connotation.

Experiment 2: Left-right auditory judgments

We next turned to the investigation of the hypothesis that the temporal meaning of binaurally presented stimulus words would influence auditory judgments. Furthermore, we examined whether the pattern of spatial distribution obtained in the first experiment would show significant convergence with the auditory judgments in the current experiment. The participants’ task was to judge whether words presented over headphones were louder on the left or the right channel. The current study avoids effects due to an overlap in response codes (Proctor & Cho, 2006), by measuring auditory left-right judgments instead of speeded left-right categorizations. The words either referred to the past or the future, or were orthogonal in meaning to time (e.g., table). We predicted
that on trials where the time-related words were presented equally loud on both channels, participants would show a judgment bias by indicating future words to be louder on the right channel more often than words referring to the past. The neutral words provided us with a baseline and were predicted to fall in the middle of past and future words.

In order to control for unequal hearing strength between the left and right ears, we calibrated the auditory stimuli for each individual separately. The response keys were aligned vertically (the “T” and “V” keys) instead of horizontally, which has been shown to activate the left-right representation of time in space (Ishihara, 2008; Torralbo et al., 2006). We did not expect the judgment bias to be completely symmetrical. As a consequence of the hemispheric asymmetry in language processing, with Broca’s speech area located in the left hemisphere, (Belin et al., 1998; Pujol, Deus, Losilla, & Capdevila, 1999; Toga & Thompson, 2003), verbal information presented to the right ear has been shown to be processed more efficiently than verbal information presented to the left ear (Belin et al., 1998; Kimura, 1961). Thus, we expected an overall shift towards a right channel disambiguation for all words. Nevertheless, a comparative analysis between the spatial positioning obtained for past and future-related words in Experiment 1 and the auditory judgments for these words in Experiment 2 was expected to show a systematic continuous overlap.

Method

Participants. Thirty-two students at Utrecht University (22 females, mean age 22) who did not participate in Experiment 1 took part in this experiment for payment or partial course credit.

Stimuli. The stimuli were identical to Experiment 1, except for the inclusion of
eight temporally neutral words (*identical, closet, even, sandal, paper, glass, table, triangle*). Neutral a-temporal words were chosen (1) to reduce the salience of the temporal dimension of the stimuli and (2) to establish a baseline against which to judge biases in the disambiguation of temporal words. Eight past and eight future-related words from Experiment 1 and the neutral words were transformed into audio files using a text-to-speech program. An analysis of the auditory stimuli with Praat speech analysis software (www.praat.org) revealed no differences in pitch or duration between the past and future words \((F’s < 1)\). Ten auditory stimuli with five sound volumes (50%, 40%, 30%, 20%, or 10% louder on the left or right channel) were created for each stimulus word. Additionally, there was a version with equal loudness on both auditory channels, which constituted the critical experimental trials.

*Procedure.* Participants were seated in individual cubicles and wore headphones. The study started with a hearing test to reduce individual differences in left-right hearing imbalances. Participants received a 440 Hz tone over headphones. If the tone was less loud on one of the channels then participants adjusted the sound until the tone was perceived to be equally loud on both channels. This procedure was repeated 3 times. The average individual adjustment was calculated, and a tone, adjusted for individual differences in left-right hearing, was played over the headphones. If participants judged the tone to be equally loud in both ears, they could continue, if not, they could redo the individual hearing adjustment. All auditory stimuli were then presented with the corresponding individual adjustment.

Participants were then instructed to listen to computer generated words presented over headphones and asked to indicate for each word whether the word they heard was
louder on the left (e.g., by pressing the “T” key) or the right channel (e.g., by pressing the “V” key). Key assignment was counterbalanced across participants.

Participants were told the meaning of the words was irrelevant for the task. They were also informed that the differences in volume between the left and right ear could be very subtle, and they should try their best to answer accurately. Each of the past, neutral, and future-related words was presented three times with equal volume on both channels (the 72 critical trials), randomly presented together with 72 filler trials. The words in the filler trials were randomly presented either 10%, 20%, 30%, 40%, or 50% louder in the left or the right ear, with each volume setting occurring twice, except for the 40% and 50% volume settings, which occurred 3 times.

Results

Since left and right auditory judgments are mutually dependent, we calculated the average percentage of times each past, neutral, and future-related word (presented three times during critical trials) was judged to be louder in the right ear, with 50% indicating an equal number of left and right channel judgments, and 100% indicating only right ear judgments. This analysis over items allows for a direct comparison of the spatial position of the temporal words in Experiment 1 with the auditory judgments in the current experiment. Response key assignment did not influence the results (all p > .80).

As expected, the average right ear disambiguation for the eight future, neutral and past-referent words differed based on their temporal meaning. The predicted linear trend (a graded right channel bias of future, to neutral to past words) was significant, $F(1, 21) = 4.72; p < .05, \eta^2_p = .26$. In support of the hypothesis under investigation, simple effects revealed that participants were more likely to judge future related words to be louder on
the right \((M = 57.16, SD = 4.67)\) than past related words \((M = 51.69, SD = 5.22)\), \(t(14) = 2.21, p < .05\). Additional analyses revealed that judgments for neutral words \((M = 55.73, SD = 5.19)\) did not differ statistically from average auditory judgments for past or future related stimuli \((p’ s > .10)\), which has to be interpreted in light of the continuous distribution of the past and future related stimuli, which consisted of time-related words ranging from extreme values (e.g., past, future) to values that were immediately adjacent to the present (e.g., a moment ago, immediately).

In line with earlier findings related to the hemispheric asymmetry in language processing (Belin et al., 1998; Kimura, 1961), there was an overall preference for right-ear judgments, which deviated significantly from guessing average, \((M = 1.65, SD = 0.16), t(23) = 4.44, p < .001\).

Finally, the hypothesized cross-modal overlap between the horizontal spatial positions of the time-referent words in the visual task of Experiment 1 with the average right ear disambiguation for each time-referent word in the auditory task in Experiment 2 was tested in a regression analysis. As predicted, the spatial position of the stimuli in Experiment 1 significantly predicted the auditory right ear judgment bias in Experiment 2, \(\beta = .55, r^2 = .30, t = 2.59, p = .03\). This linear relationship was not further moderated by the dichotomous temporal category (past vs. future, dummy coded with past as the reference group), \(\beta = 1.01, t = -0.87, p = .40\). As shown in Figure 1, the more participants placed time-referent words to the right in Experiment 1, the more often these words were judged to be louder in the right channel by participants in Experiment 2. The spatial mapping of time in the visual and auditory tasks showed a substantial cross-modal overlap.
Discussion

The current findings reveal that the temporal meaning of words influenced auditory judgments: Future-related words presented equally loud to the left and right ear were judged to be louder in the right ear more often than past-related words. Furthermore, there was a clear overlap between the spatial position of time-related words in Experiment 1 and the auditory judgments in Experiment 2. This significant cross-modal overlap underlines the argument under examination, namely that the spatial structuring of time leads to converging judgment biases in the visual and the auditory domain. Whereas previous studies have revealed that early and late auditory cues were related to left and right responses, respectively (Ishihara et al., 2008), or that auditorily presented past and future words facilitate left and right key responses (Ouellet et al., in press), this study is the first to reveal an influence of the temporal meaning of words on left and right auditory judgments.

There are a number of features of the second experiment that underline the strength of the findings and conclusions we report here. First, participants were not explicitly instructed to categorize past and future-related words (e.g., Ouellet et al., in press; Torralbo et al., 2006), or to perform explicit duration judgments (Xuan et al., 2007), but had to decide on which channel they thought each word was presented the loudest. Second, the response keys were vertically aligned, orthogonal to the horizontal dimension on which the time-related words were ordered in Experiment 1, thereby controlling for time-response congruity effects which have been observed for horizontal (but not vertical) key assignments (Ishihara et al., 2008). Third, given that correlations between explicit and implicit measures used in social cognition research tend to be quite low in
general (e.g., Fazio & Olson, 2003), we believe the current finding which reveals that the explicit visual-spatial structuring of time converges with the implicit auditory-spatial judgment bias in Experiment 2 is an important contribution. A possible avenue for future research would be to examine whether implicit measures of the spatial structuring of time, such as metaphor congruent memory biases for the spatial location of stimuli (Crawford, Margolies, Drake, & Murphy, 2006) similarly correspond with the auditory judgment bias in Experiment 2. Finally, the studies we report provide support for the assumption that the left-to-right visual and auditory structuring of temporal words is a graded continuum driven by the semantic features of the time-referent words and not a dichotomous left-right categorization.

Previous studies have revealed a dichotomous spatial representation of temporal differences when using stimulus-response interference paradigms, with the past on the left and the future on the right (Torralbo et al., 2006; Ouellet et al., 2010, in press; Vallesi et al., 2008; Weger & Pratt, 2008). A possible reason for such differences might be found by taking task demands into consideration. Speeded bimanual categorization tasks by definition require an explicit dichotomous left vs. right classification of past vs. future related stimuli (cf. Vallesi et al., 2008), which might have not been sensitive to the continuous nature of the spatial structuring of time. In the current experiments, participants were not required to perform speeded categorical responses, but performed judgments under uncertainty, which avoids structural overlap in stimulus-response mappings from influencing the results (Proctor & Cho, 2006). By revealing a spatial bias in auditory judgments in a paradigm that does not rely on response times, the current findings extend earlier results based on response interference effects in reaction time.
These findings introduce a novel multimodal methodology to investigate how abstract dimensions are structured in space, by comparing participants’ responses in a visual and auditory judgment task. Although the cross-modal convergence between the visual and auditory left-to-right representation of the abstract dimension of time provides a new contribution to the literature, previous studies have revealed cross-modal associations across many domains. For example, in studies using explicit cross-modality matching tasks, participants match bright lights with loud sounds, and dim lights with no sounds (Marks, 1978). Although our research takes a novel approach to the cross-modal overlap of time and space, the current findings have parallels with a larger body of research linking differences in horizontal space, brightness, pitch, size, and time (e.g., Dormal et al., 2006; Cohen Kadosh & Henik, 2006; Fias, Lammertyn, Reynvoet, Dupont, & Orban, 2003; Rusconi, Kwan, Giordano, Umilta, & Butterworth, 2006; Xuan et al., 2007). Whether people think about concrete perceptual dimensions such as space, or abstract conceptual dimensions such as time, certain invariant properties are extracted that share a common structure across modalities and domains.

The auditory judgment task is likely to prove to be a fertile method to investigate the cross-modal overlap in the representation of a variety of different abstract constructs. For example, given the vertical representation of power (Schubert, 2005), one would predict that powerful words presented equally loud in loudspeakers above and below the ears of a participant, are more often judged to be louder in the higher positioned loudspeaker. The findings we report here might be surprising from an amodal perspective on cognition, but follow from the grounded view on cognition that has emerged over
recent years (Barsalou, 2008; Semin & Smith, 2008; Zwaan, 2004). Our results lend further support to the fruitfulness of the assumption that abstract domains are structured in concrete sensory experiences (e.g., Boroditsky, 2000; Lakoff & Johnson, 1999). Although the possibility remains that the structuring of time in perceptual dimensions relies on purely symbolic operations, (for a more general discussion of this topic, see Barsalou, 2003, 2008; Cohen Kadosh & Walsh, 2009; Zwaan, 2009), the observed overlap between the spatial structuring of time and auditory judgments for past and future words suggests that the abstract concept of time is represented in a similar way across concrete dimensions. These insights provide interesting avenues for future research and contribute to a better understanding of how people think about abstract concepts.
References


of chronometric, neuroimaging, developmental and comparative studies of

Cohen Kadosh, R., & Walsh, V. (2009). Numerical representation in the parietal lobes:


*Cognition, 110*, 412-431.

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research:


Footnotes

1 An analysis over participants (instead of items) revealed identical results. A repeated measures analysis on the average right ear judgments for the 24 past, neutral and future words revealed an effect of words meaning $F(2, 62) = 2.77, p = .07, \eta_p^2 = .08$. A linear trend best fitted the data, with the number of right judgments being highest for future words ($M = 13.72, SD = 5.43$) and lowest for past related words ($M = 12.41, SD = 4.82$), with the neutral words falling in the middle ($M = 13.38, SD = 4.62$), $F(2,62) = 4.87, p = .035, \eta_p^2 = .14$. These analyses show that these findings can be generalized not only across words, but also across participants.
Author Note

This research was supported by the Royal Netherlands Academy of Arts and Sciences (Grant ISK/4583/PAH, awarded to the second author), and by the Fundação para a Ciência e Tecnologia (Grant PTDC/PSI/PSO/099346/2008, awarded to the third author).

Daniël Lakens is now working at Human Technology Interaction Group, IPO 1.24, PO Box 513, 5600MB Eindhoven, The Netherlands. E-mail: D.Lakens@tue.nl.
Figure Captions

Figure 1. Average horizontal position of the temporal stimuli in Experiment 1 plotted against their average percentage of right ear judgments in Experiment 2.