The Effects of Duration Words and Spatial-Temporal Metaphors on Perceived Duration

Demet Özer (dozer@ku.edu.tr)
Department of Psychology, Koç University
Rumelifeneri Yolu Sariyer 34450, Istanbul – TURKEY

Alexander Kranjec (kranjeca@duq.edu)
Department of Psychology, Duquesne University
600 Forbes Avenue Pittsburgh, PA 15282

Fuat Balcı (fbalci@ku.edu.tr)
Tilbe Göksun (tgoksun@ku.edu.tr)
Department of Psychology, Koç University
Rumelifeneri Yolu Sariyer 34450, Istanbul – TURKEY

Abstract

Subjective duration estimates are positively related to the magnitude of various non-temporal stimuli (e.g. Xuan et al., 2007). Our study investigated whether temporal and spatial magnitude information conveyed by linguistic stimuli would affect perceived duration in a temporal reproduction task. We used time-related words referring to different exact durations (e.g. *second;* Experiment 1), and spatial-temporal metaphors (e.g. *long*), referring to indistinct temporal as well as spatial magnitudes (Experiment 2). In both experiments, participants over-reproduced the shorter target duration (2.4 s) and under-reproduced the longer target duration (4.8 s). In Experiment 1, participants under-reproduced the longer target duration more when they saw "*week"* in the training and "*year"* in the reproduction. Yet, we did not observe the same semantic magnitude effect in other word pairs either in Experiment 1 or 2. Overall, we did not find supporting evidence for magnitude information conveyed by language affecting subjective time estimates.

Keywords: time perception; language

Introduction

The perception of time is a key feature of many biological and behavioral processes. Although accurate timing is essential to many daily tasks, substantial evidence shows that the subjective experience of time is not perfectly isomorphic to physical time (Zakay, 1993). Rather, perceived durations are contracted or dilated depending on many factors, including changes in non-temporal stimulus properties.

In this study, we investigated how perceived durations are modulated by temporal magnitude information provided in the medium of language.

The Interaction Between Non-Temporal Stimulus Magnitude and Perceived Duration

Subjective duration estimates are positively related to the magnitude of various non-temporal stimuli presented in different modalities. In visual domain, duration judgments were observed to be longer for larger numbers (Xuan et al., 2007; Oliveri et al., 2008; Vicario, 2011), stimulus size (Ono & Kawahara, 2007), stimulus luminosity (Goldstone, Lhamon & Sechzer, 1978), and complexity (Schiffman & Bobko, 1974). For example, people were more accurate and faster when classifying the duration of smaller magnitude numbers presented for a shorter time (congruent trials) than smaller magnitude numbers presented for a longer duration (incongruent trials) in a Stroop-like paradigm (Xuan et al., 2007). The effect of stimulus size on perceived duration has also been documented in many studies (e.g.,Ono & Kawahara, 2007; Xuan et al., 2007; Rammsayer & Verner, 2014). For example; when categorizing the durations of stimuli of different sizes by pressing one of four keys ("1" for short and "4" for long) in a temporal categorization task, people perceived larger visual stimuli as lasting longer compared to smaller visual stimuli of an equivalent duration (Ono & Kawahara, 2007).

Although the studies cited above each investigate the effects of non-temporal magnitude information on perceived duration, to our knowledge, no study so far has investigated the effects of the temporal magnitude (i.e. duration) implied by word stimuli on time perception. If there is an effect of magnitude on subjective time estimations, then we should be able to see the same effect of magnitude information derived from the semantic representations activated by linguistic stimuli. However, how semantic representations of duration and magnitude information encoded by individual words interact with the representation of duration is mostly unknown.

Interaction Between Language Processing and Low-Level Sensory/Perceptual Processing

A growing body of research investigating the interaction between language and perceptual processing suggests that semantic representations activated as we process linguistic stimuli affect the content-specific domain of low-level sensory and perceptual processing (Glenberg, Kaschak, 2002; Zwaan, 2004; Kaschak et al., 2005). According to theories of embodied language processing, comprehension involves the perceptual and motor simulation of the situation described in the linguistic input. Thus, the comprehension of words referring to a particular modal event should interact with low-level perceptual processing of that event (Barsalou, 1999; Glenberg & Kaschak, 2002; Zwaan, 2004; Kaschak et al., 2005). Many behavioral studies have provided evidence for an interaction between comprehension and perceptual processing, suggesting that higher-level semantic knowledge influences low-level sensory processing in visual perception (e.g. Spivey et al., 2001; Stanfield & Zwaan, 2001; Zwaan, Stanfield & Yaxley, 2002).

While the effects of language-activated semantic information on cognitive processing across a range of domains has been investigated, the effects of temporal magnitude representations activated by duration words and metaphors on the content-specific area of perceptual processing, namely duration perception, has not been studied. The present study aimed to fill this gap, in order to provide evidence informing both the duration perception and language processing literatures. To this end, in two experiments, we investigated how participants' reproduced duration estimations of a target interval are modulated when presented as different word types: 1) distinct temporal magnitudes (i.e. duration words; e.g. *week* vs. *year*) or 2) indistinct magnitudes or durations (i.e. spatial-temporal metaphors; e.g. *long* vs. *short*).

Experiment 1

In Experiment 1, we investigated how words referring to different exact durations (e.g. *second, year*) affect duration estimations. We hypothesized that when the word in training refers to a shorter duration compared to the word presented in the reproduction (e.g. seeing the word *"second"* in the training and *"minute"* in the reproduction), participants would under-estimate (i.e. over-reproduce) the target interval and vice versa. We did not expect any systematic difference in reproduced duration estimations when participants are presented with the same words in both reproduction and training.

Method

Participants

Twenty-five Koç University students (16 females, M_{age}=21.7) agreed to participate in exchange for course credit. We discarded one female subject because her average coefficient of variation (CV) was high (average CV across conditions = .51). All experiments were approved by the Institutional Review Panel for Human Subjects of Koç University.

Task and Stimuli

We used a temporal reproduction task. In this task, we asked participants to reproduce a given target duration by pressing a pre-designated response button to approximate the target duration as closely as possible.

At the beginning of a trial, a word ("training word") was visually presented for one of two different target intervals (2400 ms or 4800 ms). At the end of the target interval, a blank screen was presented for 1 second, followed by a fixation cross presented for a random interval between 500 ms and 1500 ms. Participants were then instructed to initiate the reproduction interval by pressing the space bar. Upon pressing the spacebar, another word ("reproduction word") appeared at the center of the screen, remaining for the entirety of the reproduction interval. The interval ended when the reproduction was perceived as temporally equivalent to the target and the participant released the spacebar. Following the termination of the reproduction interval, the next trial was presented after a random interval between 1000 and 2000 ms.

In Experiment 1, we chose four words referring to different exact durations in Turkish: *saniye ("second"), dakika ("minute"), hafta ("week") and sene ("year").* There were two conditions presented in two different sessions. In Condition 1, the words appearing in training and reproduction were different (different word pairs). In this condition, we created two-word pairs out of these four words: *("second") vs. ("minute")* and *("week") vs. ("year")*. The order of the words also changed. Thus, in some trials participants saw the word referring to the shorter duration during training (e.g. *"week"* in training and *"year"* in reproduction) and vice versa (e.g. *"year"* in training and *"week"* in reproduction), making four different training word - reproduction word pairs. In Condition 2, however, the same word appeared both in the training and the reproduction intervals (same word pairs). Thus, in this condition, four words appeared both in training and reproduction (e.g. *"week"* in training and *"week"* in reproduction, etc.).

Procedure

All words were presented at the center of the screen, printed in white on a black background. There were 30 presentations for each training word-reproduction word pair at each of the target durations. Hence, in each session, for two target durations and four word pairs, there were 240 experimental trials. We also added 24 trials (10% of the experimental trials) in which the target words appeared for a random interval between 500 and 5000 ms. We added them in order to avoid participants to habituate the two target intervals and label them as "short" and "long" durations throughout the experiment. the selected target intervals as "short" and "long" durations. Thus, in each session, there were 264 trials in total, 240 of which were used in the analyses. All trials were presented randomly. Additionally, to verify that participants looked at the screen, we asked them to report the last word they saw on the screen on 12 randomly selected trials. Participants who could not correctly identify the words three or more times were discarded from analyses (Only one participant in Experiment 2 was discarded on this front). Each participant completed the two sessions and the order of the sessions was counterbalanced across participants. Each experimental session lasted 50-60 minutes and was separated by a minimum of 1 and a maximum of 5 days.

Results

For every participant, we calculated the normalized reproduced time (i.e. the reproduced duration divided by the target duration) and averaged those scores for each word pair-target duration combination. Also, for each participant, we calculated the coefficient of variation (CV; i.e. standard deviation of each condition divided by its mean) for each condition. Reproduced intervals that were greater than three times, or smaller than one third of the target duration were excluded from the analyses. Also, the mean normalized reproduction scores and CVs that were above and below three standard deviations of the sample mean for any of the word pairs for a specific target duration were treated as outliers and excluded from further analyses.

The mean normalized reproduction times across target durations for second vs. minute and week vs. year can be found in Table 1. This table presents the over-reproduction of the 2.4 s duration and an under-reproduction of 4.8 s duration regardless of the word pair type (same vs. different) or the specific word pairs used.

Table 1: Mean normalized reproduction scores for two word pairs across two target durations. The first word in the pair is the one that was presented in training, and the second one during reproduction. The values in parentheses are the standard errors of the mean.

	2.4 s	***4.8 s***
second - minute	1.19 (.05)	.84 (.03)
minute - second	1.21 (.06)	.84 (.03)
second – second	1.22 (.05)	.89 (.03)
minute – minute	1.22 (.06)	.86 (.03)
	2.4 s	***4.8 s***
week – year	1.20 (.05)	.62 (.03)
year – week	1.15 (.04)	.83 (.04)
week – week	1.20 (.05)	.88 (.03)
year - year	1.18 (.05)	.87 (.04)

There was no effect of sex (all *p*s >.19) and the order of the conditions (all *p*s >.18) in any of the word pair–target duration combinations. Also, there was no interaction between sex and the session order (all *p*s > .22).

In Condition 1, we conducted three-way repeated measures ANOVA. The results showed that word pairs, *F* (1, 20) = 116.56, *p* < .001, target duration, *F* (1, 20) = 110.72, *p* <. 001, and the order of the word referring to the shorter duration, *F* (1, 20) = 40.60, *p* < .001, had significant main effects. However, these main effects were qualified by an interaction between all three repeated factors, *F* (1, 20) = 49.66, *p* < .001. Further comparisons showed that, for week vs. year, reproduced durations were greater when *week* was given in reproduction (*M* = .80) compared to training (*M* = .60) only when they were presented for 4.8 s.

In Condition 2, we conducted a two-way repeated measures ANOVA and found a main effect of the target duration, *F* (1, 21) = 73.35, *p* < .001, and the same word pairs, *F* (3, 63) = 4.38, *p* = .007. Pairwise comparisons showed that the mean normalized reproduced durations were greater for 2.4 s for all same word pairs compared to 4.8 s (M_{diff} = .34, *p* < .001). However, there were no significant differences between any of the same word pairs when we consider Bonferroni adjusted alpha levels of .008 per test (.05/6) in the pairwise comparisons. There was also no interaction between target duration and same word pairs, *F* (3, 63) = 1.50, *p* = .22.

To investigate the difference between the same and different word pairs, we averaged the mean normalized scores for the same and different word pairs separately and conducted a two-way repeated measures ANOVA. The results showed only a main effect of target duration, *F* (1, 19) = 92.18, *p* < .001. There was no difference between same and different word pairs, *F* (1, 19) = 3.09, *p* = .095 or the interaction between word pair type and target duration, *F* (1, 19) = 3.26, *p* = .087.

A two one-way ANOVA with all word pairs regardless of the word pair type (same vs. different) and the order of the shorter duration as repeated measures and CV scores as dependent measure was conducted separately for each target duration. For 2.4s, there was no significant effect of word pair on CV scores, *F* (4.14, 95.18) = 1.30, *p* = .25. However, for 4.8s, there was a main effect of word pair, F (7, 161) = 78.65, *p* < .001. Pairwise comparisons revealed that CVs were greater when participants saw "*week"* in the training and *"year"* in the reproduction compared to all other word pairs in 4.8 s (all M_{diff} > .173, all *p*s < .001). To see whether variability in perceived durations differed between Condition 1 and 2, we computed grand total CVs for same and different word pairs separately for each target duration and conducted a two-way repeated measures ANOVA. Results revealed a main effect of the target duration, *F* (1, 19) = 92.18, *p* < .001. The CVs were greater for 2.4 s (*M* = 1.172) compared to 4.8 s (*M* = .806). There was no main effect of the word pair type (same vs. different), *F* (1, 19) = 3.09, *p* = .095 and no interaction between two, *F* (1, 19) = 3.26, *p* =.087.

In sum, in Experiment 1, we found that regardless of the word pair type and specific order, participants over-reproduced the target duration of 2.4s and under-reproduced 4.8s. We also found that participants under-reproduced 4.8s more when they saw *"week"* in the training and *"year"* in the reproduction compared to all word pair conditions. The

CV was also greater for that word pair *("week-year")* compared to all other.

Experiment 2

In Experiment 2, we investigated how words implying both a temporal magnitude as well as a spatial magnitude modulated duration estimations. To this end, we employed quantifiers that can refer metaphorically to different indistinct durations as well as spatial magnitudes (i.e. the size of an object; e.g. *long vs. short*). Our hypotheses were same as with Experiment 1.

Method

Participants

Twenty-five Koç University students (14 females, M_{age}=21) agreed to participate in exchange for course credit. One male participant was discarded because he did not pay attention to the experiment and one female participant was discarded because her mean normalized reproduced scores were outliers in 10 out of the 16 conditions.

Task, Procedure & Stimuli

The task and the procedure were identical to Experiment 1, except for the word stimuli used in the task. In Experiment 2, we used spatial adjectives and adverbs that are used as spatial-temporal metaphors referring to indistinct durations. We chose four words: *uzun ("long"), kısa ("short"), geniş ("wide") and dar ("narrow").* In Condition 1, participants were trained with a spatial-temporal adjective and presented with the antonym of that word in the reproduction phase. We created 2 word pairs: *"long" vs. "short"* and *"wide" vs. "narrow"*. The order of the words was reversed in this condition. In Condition 2, participants saw the same spatial-temporal adjective both in training and reproduction.

Results

We used the same exclusion criteria as in Experiment 1. The averaged mean normalized reproduced intervals for short vs. long and narrow vs. wide for each target interval can be found in Table 2. Visual inspection of Table 2 suggests the over-reproduction of 2.4 s and an under-reproduction of 4.8 s in both word pairs.

Table 2: Mean normalized reproduction scores for two word pairs across two target durations. The values in parentheses are the standard errors of the mean.

	2.4 s	***4.8 s***
short - long	1.30 (.05)	.85 (.02)
long - short	1.30 (.05)	.86 (.02)
short – short	1.18 (.04)	.83 (.03)
long - long	1.20 (.04)	.83 (.03)
narrow – wide	1.30 (.05)	.88 (.02)
wide – narrow	1.27 (.04)	.86 (.02)
narrow- narrow	1.21 (.04)	.81 (.03)
wide - wide	1.23 (.05)	.82 (.03)

Neither sex (all *p*s > .12) nor the order in which participants attended the two conditions (all *p*s > .09) affected the normalized reproduction scores in any of the word pair-target duration combinations. Also, there was no interaction between sex and the session order in any of the conditions (all *p*s > .15).

For Condition 1, we conducted a three-way repeated measures ANOVA. The results revealed only a main effect of target duration, F (1, 19) = 122.96, p < .001. The mean normalized reproduction scores were greater in 2.4 s (M = 1.29) compared to 4.8 s (M = .86) for all different word pairs. There were no main effects of the specific word pair (short vs. long and narrow vs. wide), F (1, 19) = .286, p = .599, or the order of the shorter duration, F (1, 19) = .147, p = .706. Also, there was no interaction between all three repeated factors, F (1, 19) = .013, p = .910.

For Condition 2, our analysis revealed only a main effect of target duration, F (1, 20) = 69.56, p < .001. Pairwise comparisons showed that the mean normalized reproduced durations were greater in 2.4s (M = 1.18) compared to 4.8s (M = .81). There was no main effect of the same word pairs, F (3, 60) = 2.095, p = .110. There was also no significant interaction between same word pairs and target duration, F (3, 60) = 2.39, p = .078.

We conducted two separate two-way repeated measures ANOVA with the averaged mean normalized reproduced durations for the same and different word pairs for each target duration. The results showed only a main effect of target duration, F (1, 17) = 87.54, p < .001. There was no significant difference between the averaged normalized reproduced durations for same and different word pairs, F (1, 17) = 2.70, p = .119. Also, there was no interaction between two repeated factors, F (1, 17) = .999, p = .334.

With participants' CV scores, we conducted a two-way repeated measures ANOVA with all word pairs and the target duration as the two repeated factors and the CVs as the dependent measure. The results showed a significant effect of target duration, F (1, 22) = 35.338, p < .001. Pairwise comparisons revealed that CVs were greater in 2.4s (M = .260) compared to 4.8s (M = .213). There was no difference between any word pair, F (3.31, 72.83) = .639, p = .607. However, these results were qualified by an interaction between two, F (7, 154) = 2.674, p = .012. The follow-up multiple t-tests show that, when we consider Bonferroni adjusted alpha levels (.05/8 =.0062), CV scores were greater in 2.4s in word pairs "*long-short*" (M = .271), "*wide-narrow*" (M = .276), "*short-long*" (M = .259), "*narrow-wide*" (M = .265), "*wide-wide*" (M = .252) and "*narrow-narrow*" (M = .255) compared to the target duration of 4.8s (*M* = .207, .204, .201, .211, .208, .218; respectively).

In sum, in Experiment 2, we found an over-reproduction of 2.4 s and an under-reproduction of 4.8 s regardless of the word pair conditions, as in Experiment 1. However, we did not find any difference in mean normalized reproduced duration between any of the word pairs. We also found that CVs were greater in 2.4s compared to 4.8s for all four different word pairs as well as two of the same word pairs (“*short” and “wide”*).

General Discussion

In this study, we asked how language affects time perception. Specifically, we investigated how the temporal magnitude (Experiment 1; duration words) and spatial-temporal magnitude (Experiment 2; spatial-temporal metaphors) implied by words influenced subjective time estimates as assessed in temporal reproduction task. We hypothesized that increasing the magnitude conveyed by words from training to reproduction would lead to the over-reproduction of the target duration, and vice versa. We found that (1) in two experiments, participants over-reproduced 2.4s and under-reproduced 4.8s, regardless of the implied temporal / spatial magnitude of words (Figure 1 and 2), (2) CVs were greater in 2.4 s compared to 4.8 s in both experiments, and (3) participants’ reproduced durations were smaller and CVs greater when they saw *“week”* in the training and *“year”* in the reproduction in 4.8s compared to all other conditions in Experiment 1. Last, (4) we did not find any systematic effect of the temporal/spatial magnitude implied by words on perceived duration in both experiments.

The over-reproduction of 2.4s and the under-reproduction of 4.8 s in our current study are in line with Vierordt’s Law (for a review see Lejeune & Wearden, 2009) and found in many timing studies in the literature across multiple timing tasks (e.g., Karşılar & Balcı, 2016). This migration effect, which is the regression of duration estimates toward the mid-range of the target duration series, is likely due to the fact that all word pair–target duration conditions were presented randomly (i.e. interleaved) rather than in blocks.

We also detected a trend that CVs were greater for 2.4s compared to 4.8s. According to Weber’s Law, although the variation of the reproduced duration increases proportionally with the to-be-timed intervals, these results might be best explained by an additive source of variability due to experimental manipulations (other than duration) in addition to the proportional one due to timing mechanism itself (e.g. generalized form of Weber’s Law).

In Experiment 1, we found that the word pair *week-year* was under-reproduced more when presented for 4.8s compared to all other word pairs. It means that participants thought of the target duration of 4.8s as shorter when *“week”* in the training was followed by a word implying a larger temporal magnitude, like *“year”*. However, we did not see the same effect in other exact duration word pairs in Experiment 1 and spatial–temporal metaphor pairs in Experiment 2. This might be due to the larger temporal magnitude difference between these two words compared to the other word pair. Furthermore, the opposite effect was not observed for the *year-week* pair suggesting an asymmetrical form of time warping (see also Karşılar & Balcı, 2016). Further investigation is needed to determine if this effect is reliable.

Overall, we could not find supporting evidence for the effect of language on time perception. Both temporal magnitude and temporal/spatial magnitude information conveyed by words did not affect perceived duration (other than the word pair of *week-year* in 4.8s). Yet, it should be noted that there is no hypothesized model for the interaction between time perception and language. Thus, the current study is an exploratory one. However, in a recent study, Bottini and Casasanto (2010) investigated the effects of implicit spatial length information encoded in different object nouns (e.g. cigarettes, clothesline, footpath) on perceived duration and found a positive effect of spatial magnitude information conveyed in linguistic medium on time perception. Object nouns with relatively shorter implicit spatial lengths (e.g. cigarette) were remembered as appearing for shorter durations compared to nouns with longer implicit spatial lengths (e.g. footpath) despite each being presented for the same amount of time. However, we did not find the same kinds of effects. It is interesting when we consider that we used direct spatial magnitude information in Experiment 2, rather than an implicit one as in Bottini and Casasanto (2010). One possibility for falling short to replicate the findings of this study might be that the previously documented effects of magnitude on time perception are only for spatial and numerical magnitude (i.e. non-temporal) and not for temporal magnitudes. In other words, those findings might be present only for cross-domain effects. In the current study, however, we tested the impact of duration magnitude on duration perception, which is a within-domain interaction. Yet, in Experiment 2, we used spatial–temporal metaphors that implied both temporal and spatial magnitudes. One reason for the null effect in this experiment concerns the everyday use of spatial –temporal metaphors. Space and time are so intertwined that spatial adjectives are commonly understood as temporal concepts, especially in the context of a time reproduction task (Lakoff & Johnson, 1980).

Another possible explanation for not finding data to support our hypothesis in both experiments, concerns the nature of our to-be-timed stimuli. Larger, more complex, and intense stimuli expand perceived duration (Eagleman, 2008). One mechanism for this effect is the modulation of attention and arousal by the non-temporal properties of the to-be-timed stimulus. For example, intense negative sounds expand subjective duration since they heighten physiological arousal (Mella et al., 2011). Also, apart from emotional valence, attentional modulation by highly dynamic stimuli might affect duration perception. For example, Karşılar and Balcı (2016) found that higher motion coherence in a highly dynamic moving dot array may capture more attention to the non-temporal properties of the stimulus at the expense of attention to the timing task itself. This may result in the over-reproduction of a target interval when the coherence level is increased from training to reproduction. However, our stimuli were not emotionally arousing nor attention capturing. Also, magnitude was not inherently perceptible in the to-be-timed stimuli, but implied by words. Concrete, visual magnitude information presented as an inherent property of the external stimuli might affect perceived duration by better directing attentional resources to stimulus properties.

Last, the task we used might not be the most sensitive for exploring the possible effects of language on perceived duration. Other tasks, like temporal bisection (Allan & Gibbon, 1991) or categorical timing (Wearden, 1992) that force participants to decide on whether the perceived target duration is shorter or longer compared to a reference interval, might better detect differences between conditions due to its specificity to perceptual time in future studies.

In sum, the current study did not support the hypothesis that temporal and spatial magnitude information conveyed by linguistic stimuli influences subjective duration estimations. Limitations of the current study and the absence of an hypothesized model to be rejected prevent strong conclusions, but higher-order linguistic representations may not reliably interact with a low-level domain like interval timing across experimental paradigms.

References

Allan, L. G., & Gibbon, J. (1991). Human bisection at the geometric mean. *Learning and Motivation*, *22*(1), 39-58.

Barsalou, L.W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577-660.

Bottini, R., & Casasanto, D. (2010). Implicit spatial length modulates time estimates, but not vice versa. In *Spatial Cognition VII* (pp. 152-162). Springer Berlin Heidelberg.

Eagleman, D. M. (2008). Human time perception and its illusions. *Current Opinion in Neurobiology*, *18*(2), 131-136.

Glenberg, A. M. & Kaschak M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, 9, 558-565.

Goldstone, S., Lhamon, W. T., & Sechzer, J. (1978). Light intensity and judged duration. *Bulletin of the Psychonomic Society* 12.1: 83-84.

Karşılar, H., & Balcı, F. (2016). Asymmetrical modulation of time perception by increase versus decrease in coherence of motion. *Attention, Perception, & Psychophysics*, *78* (8), 2690-2707.

Kaschak, M. P., Madden, C. J., Therriault, D. J., Yaxley, R. H., Aveyard, M., Blanchard, A. A. & Zwaan, R. (2005). Perception of motion affects language processing. *Cognition*, 94, B79-B89.

Lakoff, G., & Johnson, M. (1980). Conceptual metaphor in everyday language. *The Journal of Philosophy*, *77*(8), 453-486.

Lejeune, H., & Wearden, J. H. (2009). Vierordt's The Experimental Study of the Time Sense (1868) and its legacy. *European Journal of Cognitive Psychology*, *21*(6), 941-960.

Mella, N., Conty, L., & Pouthas, V. (2011). The role of physiological arousal in time perception: psychophysiological evidence from an emotion regulation paradigm. *Brain and Cognition*, *75*(2), 182-187.

Oliveri, M., Vicario, C. M., Salerno, S., Koch, G., Turriziani, P., Mangano, R., ... & Caltagirone, C. (2008). Perceiving numbers alters time perception. *Neuroscience Letters*, *438*(3), 308-311.

Ono, F., & Kawahara, J. I. (2007). The subjective size of visual stimuli affects the perceived duration of their presentation. *Perception & Psychophysics*, *69*(6), 952-957.

Rammsayer, T. H., Verner, M. (2014). The effect of nontemporal stimulus size on perceived duration as assessed by the method of reproduction. *Journal of Vision*, 14(5):17, 1-10.

Schiffman, H. R. & Bobko, D. J. (1974). Effects of stimulus complexity on the perception of brief temporal intervals. *Journal of Experimental Psychology* 103.1: 156.

Spivey, M. J., Tyler, M. J., Eberhard, K. M., & Tanenhaus, M. K. (2001). Linguistically mediated visual search. *Psychological Science*, *12*(4), 282-286.

Stanfield, R. A. & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, 12, 153-156.

Vicario, C. M. (2011). Perceiving numbers affects the subjective temporal midpoint. *Perception- London*, 40 (1), 23.

Wearden, J. H. (1992). Temporal generalization in humans. *Journal of Experimental Psychology: Animal Behavior Processes*, *18*(2), 134.

Xuan, B., Zhang, D., He, S., & Chen, X. (2007). Larger stimuli are judged to last longer. *Journal of Vision*, *7*(10), 2.

Zakay, D. (1993). Time estimation methods do they influence prospective duration estimates? *Perception – London*, *22*, 91-91.

Zwaan, R. A., Stanfield, R. A. & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science*, 13(2), 168-171.

Zwaan, R. A. (2004). The immersed experiencer: Toward an embodied theory of language comprehension. In B.H. Ross (Ed.), *Psychology of Learning and Motivation* (Vol. 44, pp. 35-62). San Diego, CA: Academic Press.