

Response direction and sentence-tense compatibility effects: An eye tracking study

Raymond B. Becker (rbecker@cit-ec.uni-bielefeld.de)

Bridgette DeCot (bdecot@cit-ec.uni-bielefeld.de)

Ernesto Guerra (ernesto.guerra@uni-bielefeld.de)

Pia Knoeferle (knoeferl@cit-ec.uni-bielefeld.de)

Cognitive Interaction Technology Excellence Cluster and Department of Linguistics, Bielefeld University,
Morgenbreede 39, 33615, Bielefeld, Germany

Rolf Zwaan (zwaan@fsw.eur.nl)

Department of Psychology, Erasmus University
Burgemeester Oudlaan 50, 3062 PA, Rotterdam, The Netherlands

Abstract

Recent evidence shows tense-response compatibility effects only when the task relates to sentence tense (Ulrich & Maienborn, 2010). In two eye-tracking experiments, we investigated tense-response compatibility effects. In our first experiment (E1, where sentence tense was relevant to the task) we found compatibility effects at the beginning of the sentence (e.g., *Yesterday* versus *Tomorrow*), which shifted to interference effects by sentence end. Overall, we also found compatibility effects in response times, replicating Ulrich and Maienborn. Both compatibility effects in Experiment 1 (E1) were stronger for low- compared to high-WM readers. In Experiment 2 (E2, where tense was irrelevant), we found compatibility effects for high-WM readers, but only in early reading measures. These results suggest that compatibility effects are weaker depending on the task, but not eliminated; an implication which may help refine a strict view of embodied cognition.

Keywords: Mental timeline, embodiment, individual differences, eye-tracking.

Introduction

Research over the last decade has continued to refine embodiment theory (Barsalou, 1999; Glenberg, 1997), and this refinement was prodded along by criticism (Machery, 2007; Mahon & Caramazza, 2008). For example, Mahon & Caramazza argued that embodiment theory could not adequately explain how JUSTICE and other abstract concepts are understood through bodily experience because they do not reliably correspond to sensory or motor information. However, conceptual metaphor theory has laid out the groundwork for how abstract concepts such as TIME are mapped onto concrete concepts such as SPACE (Lakoff & Johnson, 1980; 1999). Torralbo, Santiago, & Lupiáñez (2006) found evidence that corroborated this potential mapping mechanism. In their Experiment 1, participants saw the silhouette of a human head looking either rightward or leftward on a screen. A word with a temporal connotation in a speech bubble was presented either in front of or behind the silhouette. Participants judged whether the person represented via the silhouette was contemplating the past or the future. When a past word appeared on the left, responses were faster than when it appeared to the right; when a future word appeared on the right side, responses were faster than

when it appeared to the left (this interaction of response-location with tense has been credited to a ‘mental timeline’, i.e., the use of a spatial left-right line to represent time in our mind). These results suggest that left- and right-hand response preparation interacts with linguistic temporal cues (past and future tense respectively). Thus it appears that abstract concepts such as TIME are grounded in experiential and bodily schemas. Meanwhile, the focus of inquiry in this area has changed from *whether* grounding effects occur for abstract concepts to *how rapidly* they occur and whether they are task-dependent. In addition, the role of participants’ working memory in these kinds of congruence effects is unclear. To contribute to these research questions we examined the time course of time- response location congruence effects during sentence comprehension as (low and high working memory) participants planned a right or left hand movement in two different tasks. Below we motivate in more detail the investigation of tense-response location congruence effects are modulated by task and working memory.

Accommodating tense-response-location congruence effects

Task appears to play an important role for tense-response location congruence effects. In a recent study, compatibility effects of tense (e.g., past versus future) and left/right response locations were eliminated in a task where tense was irrelevant. When participants paid attention to sentence tense, tense-response location compatibility effects emerged. For example, participants pressed a button labeled *Past* on the left in response to a past tense sentence more quickly than when the *Past* button was on the right. A similar compatibility effect was found for future-tense sentences and right-hand responses. However, when the task was time-irrelevant (sentence-sensibility judgments), compatibility effects were eliminated, suggesting that time-response location compatibility effects occur only when people pay attention to time. If that were the case, then both embodied (e.g., Barsalou, 1999) and non-embodied accounts such as amodal symbol systems (Collins & Loftus, 1975, Collins & Quillian, 1969) could accommodate these results. Non-embodied accounts could accommodate the

results via a traditional spreading activation network composed of disembodied, or amodal, symbols. A similar reasoning has been proposed for emotion and embodiment, but it was ruled out as it was shown that compatibility effects between emotional sentences and facial expression were task-independent (Glenberg, Havas, Becker, & Rinck, 2005). Nevertheless, the lack of tense-response compatibility effects in a time-irrelevant task (Ulrich and Maienborn, 2010) left the door open for accounts via hybrid embodiment theories (Mahon & Caramazza, 2008; Louwerse & Jeuniaux, 2008) or via cross-modal integration (Kemmerer & Gonzalez-Castillo, 2010).

There are at least two key differences between hybrid embodiment accounts and views of embodiment in which mental representations are strictly composed of perceptual symbol systems (henceforth ‘strict embodiment’). First, hybrid accounts argue that the hierarchical processing of amodal symbols occur before additional top-down context from perceptual symbols (Kemmerer & Gonzalez-Castillo, 2010) and this takes more time than a direct mapping of TIME onto SPACE. Second, strict embodiment proposes automaticity (Glenberg, 1997, p4). Automaticity refers to whether the sensorimotor system is involved in processing the meaning of abstract concepts (temporal cues in a sentence) regardless of task. Strict embodiment would thus have predicted task-independent activation of the mental timeline. Ulrich & Maienborn’s results of task-dependent tense-response location compatibility effects appeared to support a hybrid view of embodiment; however, they did not explicitly address the implications of their findings for this debate. Moreover, we cannot be certain that the lack of tense-response location compatibility effects with time-irrelevant tasks is at least in parts due to the coarse-grained response time measure they used. By monitoring eye movements during reading in addition to response times at sentence end in the same tasks that they used, we can assess the time course of tense-response location compatibility effects and determine whether the null findings are due to the nature of the measure.

Working memory and embodiment

Strict embodied cognition draws on attention and memory functions (Barsalou, 1999; Glenberg, 1997). For example, Glenberg and Gallese (2010; Koziol, Budding, & Chidekel, 2011) argued that higher order processes such as executive function are for motor control, and also part of language comprehension (see Repovš & Barch, 2012 for a possible link between working memory (WM) and cerebellum function). However, the notion of working memory as an important component of theories in cognitive psychology (see Baddeley, 2012 for a review) and psycholinguistics (Huettig, Olivers, & Hartsuiker, 2011; Lewis, Vasisht, &

Van Dyke, 2006), has been studied very little by strict embodiment theorists like Glenberg or Barsalou¹.

Due to the scarcity of research on embodiment and working memory, we drew from work on temporal order and working memory (Münte, Schiltz, & Kutas, 1998). This previous research suggested that participants with high-WM used temporal cues such as *Before* versus *After* immediately to aid sentence processing. For example, in the following sentences, the initial temporal adverb and verb tell the reader that this sentence describes an event that occurred in the past (1) or the future (2):

(1) *Früher in dieser Woche*_{ADV} | *faltete*_{VP} | *Jennifer*_{NP1}
*im Wohnzimmer*_{PP} | *die Wäsche*_{NP2}.

‘Earlier this week_{ADV} | folded_{VP} | Jennifer_{NP1}
in the living room_{PP} | the laundry_{NP2}’.

(literal translation).

(2) *Später in dieser Woche*_{ADV} | *faltet*_{VP} | *Jennifer*_{NP1}
*im Wohnzimmer*_{PP} | *die Wäsche*_{NP2}.

‘Later this week_{ADV} | folds_{VP} | Jennifer_{NP1}
in the living room_{PP} | the laundry_{NP2}’.

It would be consistent with the findings of Münte, Schiltz, and Kutas (1998) if high- (but not low) WM immediately processed the temporal cue ‘Earlier / Later’. A question that could be asked with respect to the role of working memory in embodied cognition is whether participants immediately integrate temporal cues in addition to response-location as they make a sensibility judgment about a sentence.

The present study

Using eye-tracking, the present studies thus investigated the time course of tense-response location compatibility effects as a function of (a) task (time-focus vs. no time focus), and (b) participants’ working memory. The use of eye tracking and a between-experiment task manipulation permitted us to test the strict embodied hypothesis.

In Experiment 1, participants performed a tense evaluation task (was the sentence in the past or in the future)? One gaze pattern in support of strict embodiment would be an early-peaking, quickly decaying Simon-like effect consistent with the action and perception literature (Symes, Tucker, & Ellis, 2005). This pattern would suggest both a rapid (tied to first-pass measures) and automatic (insensitive to task) tense-response location congruence effects in first-pass times at the verb and potentially also the

¹ We would like to thank an anonymous reviewer for suggesting the Coherent Working Models theory proposed by Santiago and colleagues. Unfortunately, due to time and page limit constraints, we have not integrated their proposal with our current framing or the discussion of our results. However, we will review the data supporting the Coherent Working Models theory, and how well our results fit with it in an extended manuscript in preparation.

next sentence region. If it is rapid and automatic but quickly decays, then there may be no effect in a relatively late and course-grained measure such as response times. By contrast, the amodal views would predict no tense-response compatibility effect for time-irrelevant tasks: symbols representing the concepts of FUTURE and PAST would not become bound to right and left procedural symbols and in turn not become activated. Hybrid accounts, would predict task-independent effects but these should occur later than in strict embodiment accounts.

We further hypothesized that high-WM readers would rapidly process the sentence-initial temporal cue (see Münte, Schiltz, & Kutas, 1998). In order to come up with a hypothesis regarding working memory and tense-response location congruence effects, we drew from one of the more important ideas in embodied cognition which is that people can “offload” cognition to the environment (Clark, 1997; Spivey, 2008). To the extent that cognitive load can be offloaded onto the environment, a left-right mental timeline in which left indexes the past and right the future could assist in processing temporal information. And if the results from Münte et al. generalize to tense-response location congruence effects, then high-WM participants should process tense-response location congruence earlier. We predicted longer first-pass times in the subject noun phrase region (NP1 in sentences (1) and (2)) for incompatible (vs. compatible) tense-response location for high-WM participants because the region is potentially where participants would shift their attention from the tense processing to the sensibility judgment. This effect could extend to the locative prepositional phrase region as well because the attention shift could take time even for high-WM readers (see sentences (1) and (2)). By contrast, for low-WM readers congruence effects should emerge at the end of the sentence, because in Münte et al. low-WM readers did show evidence that they were processing the temporal cues but later than the high-WM readers. In response times, both groups should show a compatibility effect replicating Ulrich and Maienborn (2010).

When time is not relevant for the task (Experiment 2), and if the null effect in Ulrich and Maienborn is an artifact of the post-sentence response time measure, then we should see similar yet more subtle effects than in E1 (compatibility effects in first-pass times, potentially also earlier for high-WM than low-WM readers). However, these patterns of reading times should not result in compatibility effects in response times based on the findings by Ulrich and Maienborn.

Experiment 1: Time-relevant task

In Experiment 1, we replicated the procedure of the first experiment by Ulrich and Maienborn (2010). Participants were asked to pay attention to sentence tense and registered their decision via a button press if the sentence made sense. Thinking about time was part of the task because participants made explicit decisions about sentence tense.

Method

Participants 48 members (17 male with mean age 24; $SD = 3$ years) of Bielefeld University participated in the experiment. All participants were native German speakers with no second language exposure prior to 6 years of age; had normal or corrected-to-normal vision; were naïve with respect to the purpose of the study; and received €6 for their participation or course credit. All gave informed consent.

Materials Items consisted of 48 past and future tense sentences beginning with a temporal adverb and 48 nonsense sentences of the same syntactic structure. The nonsense sentences included the same words as those used in Ulrich and Maienborn (2010), but were restructured to be similar to the critical sentences, where the temporal adverb was always at the beginning of the sentence.

Procedure Participants were asked to judge whether the sentence referred to the past or future. However, they were also asked to respond only if the sentence made sense (see Figure 1). For nonsense sentences, participants were instructed to wait until the trial timed out. Their eye-movements were recorded at 1000 Hz using an EyeLink 1000 desktop mounted tracker. Participants’ WM was tested by the automated reading span test (Unsworth, Redick, Heitz, Broadway, & Engle, 2009).

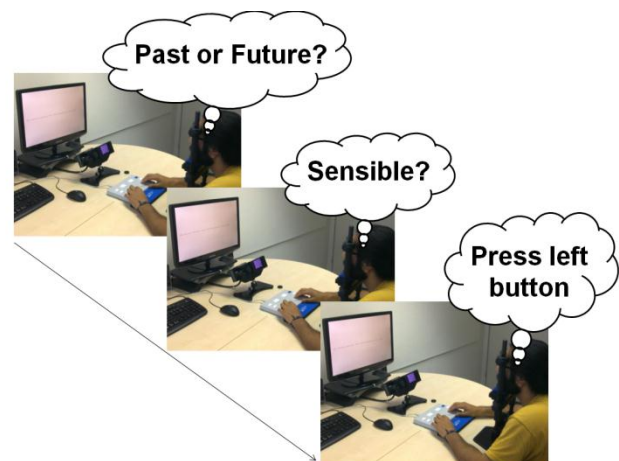


Figure 1: Example trial from E1.

Analysis

Prior to analysis fixations were cleaned using a 4-stage procedure. In the first stage fixations less than 80ms were merged with the nearest neighboring fixation if it was longer than 80ms and within 0.5 degrees of visual angle away along the x-axis. Similarly, in the second stage fixations less than 40ms were merged with the nearest neighboring fixation if it was longer than 40ms and within 1.25 degrees away along the x-axis. In stage 3, every interest area was checked for at least three fixations less than 140ms and none larger than 140ms. If an interest area was found that met these criteria, these fixations were merged with the larger ones. Lastly, all fixations less than 80ms and 1200ms were

removed. Trials with incorrect answers and nonsense sentences were not analyzed. Participants were split into high- and low-WM groups using a tertile split forming three groups of 16 people each, but only the high- and low-WM groups were included in order to do an extreme groups analysis. We conducted a 2 (WM) x 2 (tense) x 2 (response location) linear mixed effects model analysis to test for tense-response compatibility effects in each sentence region; starting with the full model and removing parameters until we found the most parsimonious model that best fit the data (Baayen, Bates, & Davidson, 2008)².

Results

There were no significant effects in first-pass reading times for any sentence region. In total dwell times in the sentence-initial temporal adverb region, we found a significant tense-response compatibility effect, $t(30) = -2.16, p < 0.05$ (see Fig. 2). Further, low-WM readers showed the compatibility effect, whereas high-WM readers did not, as evidenced by a 3-way interaction, $t(30) = 2.22, p < 0.05$ (see Fig. 3). Surprisingly, *interference* effects emerged at sentence end for both groups, $t(30) = 2.08, p < 0.05$ (see Fig. 4). In response times, we replicated the congruence effect from Ulrich & Maienborn, (2010), $t(30) = -3.87, p < 0.05$. However, the congruence effect was driven by the low-WM group as evidenced by a 3-way interaction, $t(30) = 2.71, p < 0.05$ (see Fig. 5).

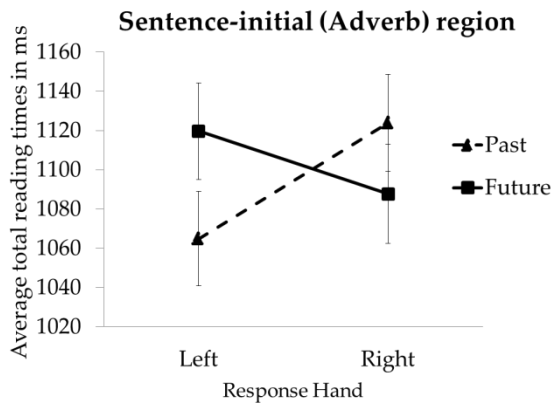


Figure 2: Tense-response location compatibility effects for both WM groups for total dwell times in the temporal adverb region. Error bars indicate the standard error (SE).

² Thank you to an anonymous reviewer for pointing out the fact that eye-tracking researchers enjoy many degrees of freedom in their research (e.g., regions of interest, first-pass readings versus total times, etc...). And further that our effects are quite small and potentially would not stand up to Bonferroni correction. Because we are in a crisis in psychology of false positives and failures to replicate, we used linear mixed effect models, backward model selection, and report the pMCMC values (Barr, Levy, Scheepers, & Tily, 2013).

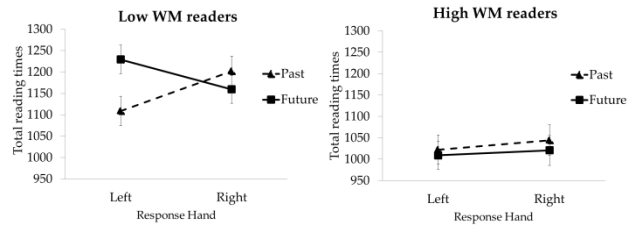


Figure 3: Tense-response location compatibility effects for low- (left) and high-WM (right) groups for total dwell times in the temporal adverb region. Error bars indicate the SE.

Discussion

The response time results from Experiment 1 replicate prior tense-response location congruence effects in response times (Experiment 1 in Ulrich and Maienborn, 2010). However, the pattern of the compatibility effects over the course of the sentence, and as a function of working memory showed that there is more to the story. The response time compatibility effects in the low-WM group, but not the high-WM, are similar to the pattern for both groups. The same pattern can also be seen in the total reading times of the sentence-initial region (e.g., ‘earlier’ / ‘later’) for low-WM readers only.

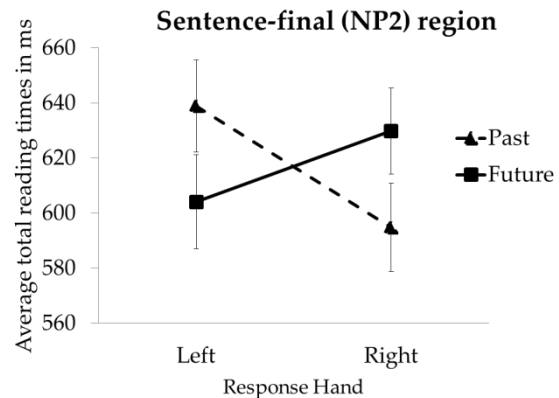


Figure 4: Tense-response location interference effects in total times in the sentence-final region. Error bars indicate the SE.

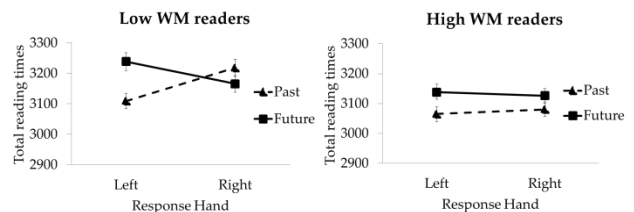


Figure 5: Tense-response location compatibility effects for low- (left) and high-WM (right) groups for response times. Error bars indicate the SE.

Next, we attempted to replicate the procedure of Ulrich and Maienborn’s (2010) Experiment 2 to assess whether we would find early compatibility effects undetectable in full-sentence response time, and how that may vary as a function of working memory ability. Participants’ task was to make a

sentence-sensibility judgment, and then press a button to indicate their decision. In this case thinking about time is *irrelevant* to the task, because it is not the decision that participants have been asked to make about the sentence.

Experiment 2: Time irrelevant task

Method

Participants We tested a further 48 students (11 male with the mean age of 23; $SD = 3$ years) who met the same criteria as those in E1.

Materials and Procedure The materials were identical to those in E1. Participants judged sentence sensibility, thus time, or tense, was irrelevant to the task (Fig. 5). This procedure is identical to the second experiment in Ulrich & Maienborn (2010). Again, the only difference was that the sentences always had the same word order and always included a prepositional phrase after the verb.

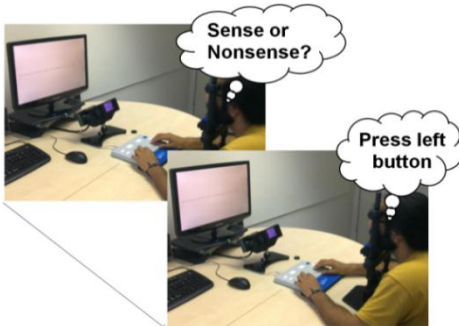


Figure 6: Example trial from Experiment 2.

Results

Data filtering and separating of participants into two WM groups was done in the same way as in E1. For first-pass reading times, we found a significant 3-way interaction in the verb region, $t(30) = -2.07, p < 0.05$ (see Fig. 7).

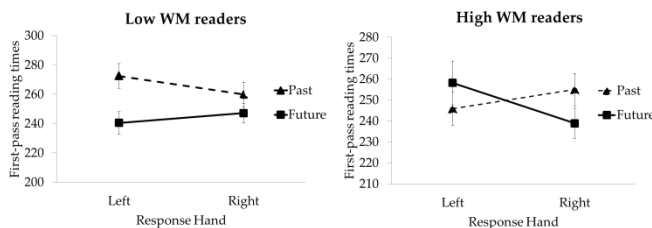


Figure 7: Sentence-tense compatibility effects for low- (left) and high-WM (right) groups for first-pass reading times in the verb region. Error bars indicate the standard error.

Next, both first-pass, $t(30) = 2.11, p < 0.05$, and regression-path duration, $t(30) = 2.15, p < 0.05$, revealed a significant interaction in the sentence-initial temporal adverb region. For both measures, durations were longer when the adverb indicated a past tense sentence and the participants were planning a left response compared to a right response, whereas for future-indicative adverbs there was no reliable difference between left and right response locations. Lastly, we replicated the absence of reliable

compatibility effects in response times for tense (see Figure 8 and Ulrich & Maienborn, 2010).

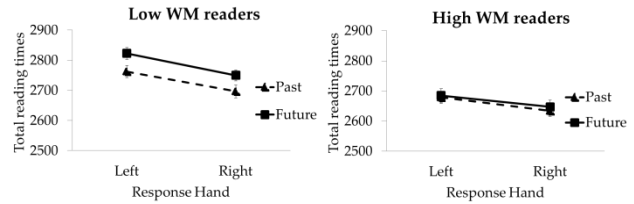


Figure 8: This figure illustrates the absence of sentence-tense compatibility effects for low- (left) and high-WM (right) groups in response times. Error bars indicate the SE .

General Discussion

Consistent with Ulrich and Maienborn (2010), we found compatibility effects in the response times when participants judged sentence tense (Experiment 1, E1) but not when they judged sentence sensibility (Experiment 2, E2). However, eye tracking revealed additional details about the time course and individual differences of the compatibility effects in E1 and the null finding in E2.

When the task was to decide on sentence tense (E1), we found compatibility effects in the earliest sentence region (e.g., ‘Earlier’), but in a relatively “late” measure; in total times. This replicated the region of the sentences where temporal processing occurred, similar to Münte et al. (1998) but it is later in the time course than we initially predicted. In contrast to Rapid Serial Visual Presentation used by Münte et al., total times include re-readings and thus potentially later processes. One possible reason why compatibility effects emerge only at this region and not in the verb region is because that region is central to judging sentence tense, and foregrounds tense processing.

Two further unexpected findings in Experiment 1 were the sentence-final interference effect and that the response time compatibility effect was driven by the low-WM readers. The interference effect could index that as participants’ prepare for the tense decision and gauge sensibility at the end of the sentence, they may momentarily inhibit temporal information. Because the temporal information has already been mapped into the environment, freeing up resources from WM needed for further language comprehension may be aided by the suppression of environmental patterns, in this case tense-response mappings (Glenberg, 1997, p4). The compatibility effects for low-WM readers at the sentence-initial temporal region are consistent with this idea: Perhaps low WM-readers are slower than high-WM readers to inhibit tense information, and thus show compatibility effect in total times at the sentence-initial region while these effects are absent for high-WM readers. The assumption here is that low-WM readers are slower because they have more difficulty updating their WM, which according to Glenberg is a conscious and effortful use of memory.

For Experiment 2, when the task did not involve a sentence-tense decision, we replicated the absence of a

tense-response location compatibility effect in response times (Ulrich & Maienborn, 2010). By contrast, compatibility effects emerged as predicted at the verb region in first-pass reading times, but only for high-WM readers. The null effect for low-WM readers may indicate that unless the temporal information is part of the task, low WM readers do not integrate tense information in relation to response location.

Overall, thus, tense-response location compatibility effects varied as a function of task and comprehenders' working memory. Our findings highlight the importance of using continuous measures: While end-of-sentence response times suggested task can eliminate compatibility effects, these effects were clearly present in gaze measures during sentence reading even when the task did not ask participants to focus on tense cues. With regard to embodiment theory, it seems that tense-response location compatibility effects are not eliminated by tasks in which tense is irrelevant but there is a need to accommodate their variation by task and working memory.

Acknowledgments

This research was funded by the Cognitive Interaction Technology Excellence Center (German Research Foundation, DFG).

References

- Baayen, R.H., Davidson, D.J. and Bates, D.M. (2008) Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390-412.
- Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, 63, 1-29.
- Barr D. J., Levy R., Scheepers C. & Tily, H. (2013). Random-effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255-278.
- Barsalou, L. W. (1999). Perceptions of perceptual symbols. *Behavioral and Brain Sciences*, 22, 637-660.
- Clark, A. (1997). *Being there: Putting brain, body, and world together again*. Cambridge, MA: The MIT Press.
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82, 407-428.
- Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 8, 241-248.
- Glenberg, A. M., Havas, D., Becker, R., & Rinck, M. (2005). Grounding Language in Bodily States: The Case for Emotion. In R. Zwaan and D. Pecher (Eds.) *The grounding of cognition: The role of perception and action in memory, language, and thinking*. Cambridge: Cambridge University Press.
- Glenberg, A. M., & Gallese, V. (2010). Action-based language: A theory of language acquisition, comprehension, and production. *Cortex*, 48, 905-922.
- Huetting, F., Olivers, C. N. L., Hartsuiker, R. J. (2011). Looking, language, and memory: Bridging research from the visual world and visual search paradigms. *Acta Psychologica*, 137, 138-150.
- Kemmerer, D., & Gonzalez-Castillo, J. (2010). The two-level theory of verb meaning: An approach to integrating the semantics of action with the mirror neuron system. *Brain and Language*, 112, 54-76.
- Koziol, L. F., Budding, D. E., & Chidekel, D. (2011). From Movement to Thought: Executive Function, Embodied Cognition, and the Cerebellum. *Cerebellum*, 11, 505-25.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.
- Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh*. New York: Basic Books.
- Lewis, R. L., Vasisht, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences*, 10, 447-454.
- Louwerse, M. M., & Jeuniaux, P. (2008). Language comprehension is both embodied and symbolic. In M. de Vega, A. Glenberg, & A. C. Graesser (Eds.), *Symbols and embodiment: Debates on meaning and cognition* (pp. 309-326). Oxford: Oxford University Press.
- Machery, E. (2007). Concept empiricism: A methodological critique. *Cognition*, 104, 19-46.
- Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology-Paris*. 102, 59-70.
- Münte, T. F., Schiltz, K., & Kutas, M. (1998). When temporal terms belie conceptual order. *Nature*, 395, 71-73.
- Repovš, G., & Barch, D. M. (2012). Working memory related brain network connectivity in individuals with schizophrenia and their siblings. *Frontiers in Human Neuroscience*, 137, 1-15.
- Spivey, M. (2008). *The continuity of the mind*. New York, NY: Oxford University Press.
- Spivey, M., & Geng, J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research*, 65, 235-241.
- Symes, E., Ellis, R., & Tucker, M. (2005). Dissociating object-based and space-based affordances. *Visual Cognition*, 12, 1337-1361.
- Torralbo, A., Santiago, J., & Lupiáñez, J. (2006). Flexible conceptual projection of time onto spatial frames of reference. *Cognitive Science*, 30, 745-757.
- Ulrich, R., & Maienborn, C. (2010). Left-right coding of past and future in language: The mental timeline during sentence processing. *Cognition*, 117, 126-138.
- Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J. M., & Engle, R. W. (2009). Complex working memory span tasks and higher-order cognition: A latent-variable analysis of the relationship between processing and storage. *Memory*, 17, 635-654.