Article

Six Challenges for Embodiment Research

Markus Ostarek[1] [iD] and Falk Huettig[1,2]

[1]Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands, and
[2]Centre for Language Studies, Radboud University

Abstract

Twenty years after Barsalou's seminal perceptual-symbols article, *embodied cognition*, the notion that cognition involves simulations of sensory, motor, or affective states, has moved from an outlandish proposal to a mainstream position adopted by many researchers in the psychological and cognitive sciences (and neurosciences). Though it has generated productive work in the cognitive sciences as a whole, it has had a particularly strong impact on research into language comprehension. The view of a mental lexicon based on symbolic word representations, which are arbitrarily linked to sensory aspects of their referents, was generally accepted since the cognitive revolution in the 1950s. This has radically changed. Given the current status of embodiment as a main theory of cognition, it is somewhat surprising that a close look at the literature reveals that the debate about the nature of the processes involved in language comprehension is far from settled, and key questions remain unanswered. We present several suggestions for a productive way forward.

Since Barsalou's (1999) seminal article on perceptual symbols, a large literature has emerged that addresses the extent to which the evolutionarily old sensory, motor, and affective systems are involved in high-level cognition (Meteyard, Cuadrado, Bahrami, & Vigliocco, 2012). Although various flavors of embodiment have been proposed (Wilson, 2002), we focus on the notion of *simulation*, the activation of sensory states that resemble those active during action and perception, as a crucial cognitive mechanism. Here, we discuss six challenges for embodiment from the perspective of language processing, but our arguments are applicable to the field as a whole.

Developing Decisive Paradigms

In order to develop good theories of embodiment, we need paradigms that directly probe the hypothesized perceptual processes. Large parts of the embodiment literature and the ensuing theoretical claims are based on congruency effects between aspects of linguistic and pictorial stimuli or actions. In the classic sentence-picture-verification (SPV) paradigm, sentences implying that an object has a certain shape lead to faster responses to pictures of objects that are congruent (vs. incongruent) with that shape (Zwaan, Stanfield, & Yaxley, 2002).

Congruency effects are important because they are straightforwardly predicted from embodied-cognition accounts but appear counterintuitive on the basis of symbolic amodal theories. Basic congruency effects, however, do not allow us to arbitrate between embodied and amodal accounts because language and picture stimuli (or actions) all involve multiple processing stages, and on the basis of accuracy rates and reaction times (RTs), it is impossible to know at what stage interactions arise (Mahon & Caramazza, 2008). An alternative to the simulation-based interpretation of the SPV results (Zwaan et al., 2002) is that comprehenders extract abstract shape information from the sentences and pictures that are represented in a symbolic conceptual module, and hence a congruency effect arises. Standard congruency paradigms are useful to establish the informational content that is activated, but they cannot determine the nature of the underlying processes.
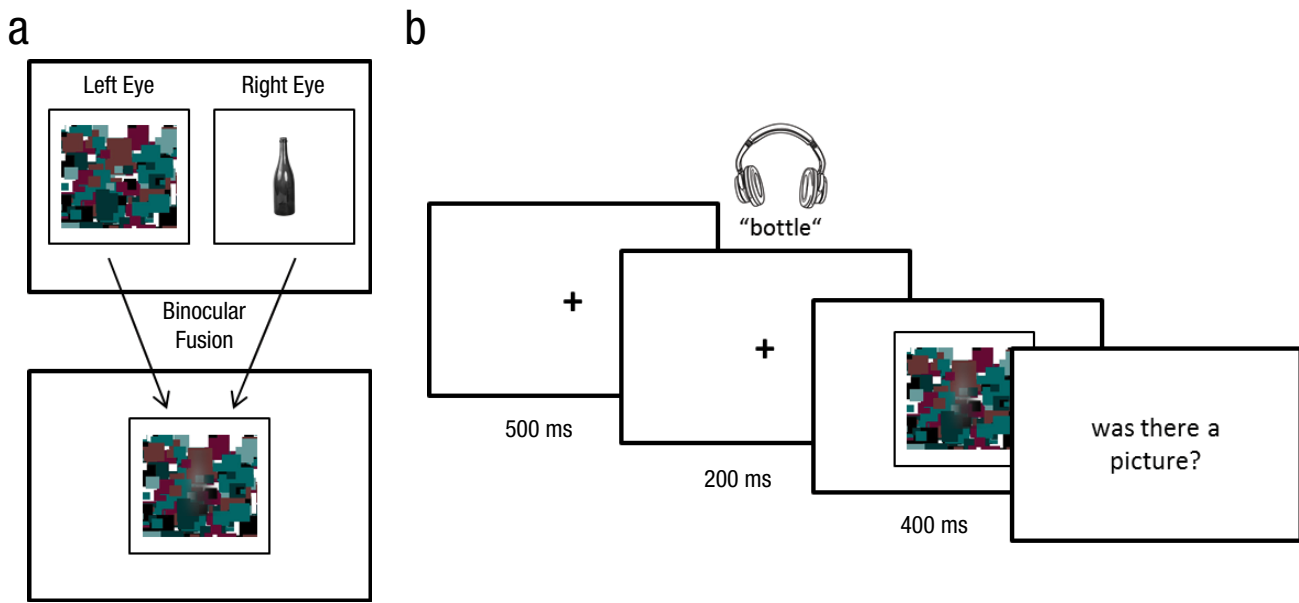
One way to circumvent the interpretational ambiguity of congruency effects is to create situations in which stimuli (that language is predicted to interact with) are

Corresponding Author:
Markus Ostarek, Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, The Netherlands
E-mail: markus.ostarek@mpi.nl

Fig. 1. Continuous flash suppression (CFS). In the CFS paradigm, a static image of an object is presented to one eye while a dynamic mask is presented to the other (a). This leads to suppression of the object, and the strength of suppression can be set to individual participants' visibility threshold. In a typical CFS trial (b), participants hear an audio cue that either describes or does not describe the suppressed image (congruent and incongruent words, respectively). This makes it possible to test the effect of processing congruent versus incongruent words on detection sensitivity. Adapted from Ostarek and Huettig (2017a).

processed only at the level of interest, for instance, the visual level. However, many seemingly perceptual tasks involve high-level categorization processes, and common masking techniques, such as backward masking, do not consistently preclude semantic processing (Kouider & Dehaene, 2007). A technique that can overcome this problem is continuous flash suppression (CFS), because detection in CFS depends on how efficiently visual features of suppressed pictures are processed (Fig. 1).

Two recent studies showed that hearing spoken words increases detection sensitivity to congruent pictures in CFS, suggesting that language modulates early stages of conscious vision (Lupyan & Ward, 2013; Ostarek & Huettig, 2017a). The finding that language can modulate detection in CFS constitutes much stronger evidence for the view that semantic processing engages perceptual processes. In short, for congruency paradigms to be informative about embodiment, great care needs to be taken to isolate the processes of interest. CFS is a promising method in this context, and we propose that it should become the benchmark test of visual simulation.

Probing Causality

Evidence for a causal link between simulation and language comprehension must be the gold standard for assessing embodied accounts. Interference paradigms are one way to probe causality. Converging evidence suggests that visual noise interferes with access to visual information during word processing, at least in some tasks: Visual noise makes word cues less effective with regard to visual (e.g., "Is it round?") but not categorical (e.g., "Is it furniture?") semantic information in word-picture verification (Edmiston & Lupyan, 2017). Further, it slows down responses to highly "imagable" concrete nouns compared with abstract nouns (Ostarek & Huettig, 2017b) in a concreteness task (in which visual information is relevant) but not in a lexical decision task (LDT) and a word-class judgment task (in which visual information is not relevant). Causal evidence for the role of simulation in language comprehension beyond single words is still missing. A recent study that tested the effect of different types of visual interference on the shape-match effect in the SPV paradigm suggests limits for a causal role (Ostarek, Joosen, Ishag, De Nijs, & Huettig, 2019), but further research is needed to determine which types of processes are active in SPV.

Transcranial magnetic stimulation (TMS) is a particularly useful neuroscientific method for probing causality in the motor domain. Given that it is easy to localize effector-specific motor areas and that TMS affords precise control of timing, surprisingly little work has been published in this area. Facilitatory effects on RTs in an LDT were obtained by using two-pulse TMS on the primary motor cortex (M1; Pulvermüller, Hauk, Nikulin,

& Ilmoniemi, 2005) and by using off-line TMS on the premotor cortex (Willems, Labruna, D'Esposito, Ivry, & Casasanto, 2011). In contrast, another study found that four-pulse TMS on M1 interfered with RTs in a concreteness task but found no robust effects in an LDT (Vukovic, Feurra, Shpektor, Myachykov, & Shtyrov, 2017). Thus, the available evidence does not paint a clear picture of the conditions in which facilitation, interference, or no effect is expected. Given the replication crisis in psychology (Zwaan, Etz, Lucas, & Donnellan, 2018)—a crisis from which some well-known embodied-cognition effects in the literature have not been exempt— a large-scale, preregistered, multilab effort applying TMS to probe causality of embodied language processing would be particularly fruitful.

Understanding the Task Dependency of Embodied Language Processing

Similarly important is the systematic study of task and context effects. The available evidence strongly suggests that semantic processing is not an automatic process (Lebois, Wilson-Mendenhall, & Barsalou, 2015). Instead, contextual factors, such as the current task requirements, can lead to surprising situations in which, for instance, the word *red* primes a green target more than a red target (Merikle, Joordens, & Stolz, 1995). Importantly, recent studies have begun to delineate the situations in which sensory processes are causally involved in language comprehension. The emerging picture is that visual and motoric processes are functionally relevant in situations in which visual and motor information is required by the task (Edmiston & Lupyan, 2017; Ostarek & Huettig, 2017b; Tomasino, Fink, Sparing, Dafotakis, & Weiss, 2008; Vukovic et al., 2017).

A crucial issue thus arises: To what extent does embodied language processing occur routinely? Different contexts make different aspects of word meanings relevant (Estes & Barsalou, 2018). It is conceivable that simulations are specifically tailored to fit those dynamic contextual demands. Nevertheless, it is apparent that most of the causal evidence stems from experiments that directly probe sensory properties of word referents. How representative are such tasks of the processes that are activated during everyday communication? A deflationary possibility is that sensory processes are important in only rare situations in which one consciously reasons about sensory aspects of word meanings. On the other hand, knowledge of perceptual information is likely required in most communication situations, as people often talk about things in the here and now. It is thus necessary for future research to go the extra mile and conduct experiments in naturalistic settings to establish the extent to which simulations are a routine part of language processing "in the wild" and the extent to which semantic context determines the deployment of simulations.

Explicit Predictions About the Direction and Timing of Effects

Lacking predictions about the direction of an effect typically reflects that the tested theory is not specific enough, leading to a situation in which opposite experimental outcomes are interpreted as evidence for the theory. For example, Connell (2007) used an SPV experiment to test whether "mentally representing something *red* engages the neural subsystems that respond to environmental perception of that colour [sic]" (p. 476) and observed that pictures that mismatched the color implied in sentences led to reduced RTs. Several other investigators obtained the more intuitive finding that a match leads to shorter RTs (e.g., Mannaert, Dijkstra, & Zwaan, 2017). All authors concluded that color is simulated during language comprehension (although see Zwaan & Pecher, 2012, for a nuanced discussion). The fact that opposing outcomes are interpreted as supporting the same claim raises serious issues about falsifiability (Popper, 2005) that can be overcome only by theories that commit to specific claims about the nature and timing of the hypothesized mechanisms and about how these play out in a given experimental paradigm.

This is not only necessary but also achievable. Many studies have investigated the effects of words with vertical spatial associations (e.g., *cloud* is up, *foot* is down) on visual detection and discrimination tasks involving targets in up or down locations (Gozli, Chasteen, & Pratt, 2013). Implicit up/down words tend to interfere with targets in compatible locations in discrimination tasks when words and targets are semantically unrelated and targets appear within 400 ms after the word, whereas facilitation is observed in detection tasks when the time between words and targets is longer than 400 ms or when semantically related words and targets are used (Estes & Barsalou, 2018; Gozli et al., 2013). A theoretical account that explains the direction and timing of effects is that words trigger visual simulations of typical events in visuospatial systems (Ostarek & Vigliocco, 2017). The theory predicts that in this paradigm, parts of mental space (top, bottom) are attended to and are temporarily occupied by simulated objects for around 400 ms (the typical semantic-processing duration for single words). Thus, interference is predicted when the simulated object is incompatible with a to-be-identified visual target, but facilitation is predicted when it is compatible. This is precisely what recent behavioral studies observed (Estes, Verges, & Adelman, 2015; Ostarek & Vigliocco, 2017). Further

evidence demonstrated that initial vertical attention shifts and facilitation or interference resulting from a match or mismatch of the simulated and observed object are reflected sequentially in saccade trajectories (Ostarek, Ishag, Joosen, & Huettig, 2018).

The Need for an All-Encompassing Theory

Embodied accounts make a good case for language processing about concrete objects and situations; however, a convincing case for how abstract concepts fit into the picture has not been made (Borghi et al., 2017; Dove, 2016). For instance, it is unlikely that a sentence such as "dignity is a basic right" will be fully captured by sensorimotor simulation, by the emotions associated with it (Kousta, Vigliocco, Vinson, Andrews, & Del Campo, 2011), or by activating a concrete domain with partially overlapping features (Lakoff & Johnson, 2003). More generally, embodiment does not provide an intuitive account of those aspects of conceptual processing that depend on high degrees of abstraction from specific events. Recognizing that simulation is not a suitable solution for certain aspects of semantic cognition, such as abstraction, thematic relationships, or fast linguistic associates, pluralistic accounts have recently been proposed in which amodal and grounded (modal) representations coexist and engage in a division of labor (e.g., Zwaan, 2014). In a particularly promising line of research, scientists have begun to unravel the hierarchical organization of conceptual content in the brain and have posited a pivotal role for multimodal abstraction in association areas (Binder, 2016; Fernandino et al., 2016; see also Barsalou, Dutriaux, & Scheepers, 2018).

This is a slippery slope, however, because the need to ground symbols in sensorimotor systems does not disappear and hide while towers of abstraction are erected. The main challenge for embodied theories of language remains to provide a coherent account of abstraction while taking sensorimotor grounding seriously. This is most likely to succeed in models whose core principles naturally afford abstraction capacities rather than models in which multiple qualitatively different systems coexist.

Assessing Emerging Theories With Novel Neuroimaging Methods

Novel neuroimaging techniques can be highly valuable as new theories emerge. A particularly promising tool is multivariate pattern analysis (MVPA), which can be exploited to ask hypothesis-driven questions about the information reflected in activation patterns. For instance,

Anderson, Bruni, Lopopolo, Poesio, and Baroni (2015) compared the abilities of a visual model (Fig. 2) and a text-based model to predict activation patterns elicited by written words and demonstrated that patterns in the visual and ventral-temporal cortex reflected visual features of the word referents that participants were thinking about. In a similar spirit, Borghesani and colleagues (2016) demonstrated that the real-world object size of word referents is reflected in activity patterns in the early visual cortex, whereas categorical semantic structure was reflected in anterior temporal regions.

Intracranial recordings deliver high temporal and spatial resolution that, in combination with MVPA, can be used to put specific hypotheses to the test. In rare cases, MVPA can be done at the single neuron level, making it possible to test whether sensorimotor neurons are also tuned to sensorimotor aspects of word meanings (Yang et al., 2017), which constitutes a direct test of the simulation hypothesis in its strongest form.

Working memory and imagery have recently been found to share activity patterns with visual processing in the visual cortex (Harrison & Tong, 2009). A recent lamina-resolved functional MRI study (Lawrence et al., 2018) suggests that bottom-up signals about the orientation of gratings predominantly reach Layer 4 of the primary visual cortex, whereas top-down imagery signals target the deep and superficial layers (Fig. 3). The dissociation between bottom-up and top-down signals was specific to the primary visual cortex; in the subsequent regions of the visual pathway, all layers were activated to the same extent. Extrapolating to simulation in language comprehension, this could be taken as an upper bound on the similarity between perception and simulation: The lowest processing stage targeted by simulation signals is expected to dissociate from sensory processing at the laminar or neural level, but activation patterns may be shared in subsequent processing stages. The types of simulated features are likely to determine which regions are targeted by simulations.

Summary

Embodiment research faces important challenges, including a shift toward more decisive paradigms, allowing the field to pinpoint the situations in which sensorimotor processes contribute functionally to language comprehension. Although improved behavioral and neuroimaging approaches can go a long way, theoretical progress will be crucial to transform embodiment from an underspecified general framework of related ideas to a fully specified complete theory of semantic cognition from which precise experimental predictions can be derived.
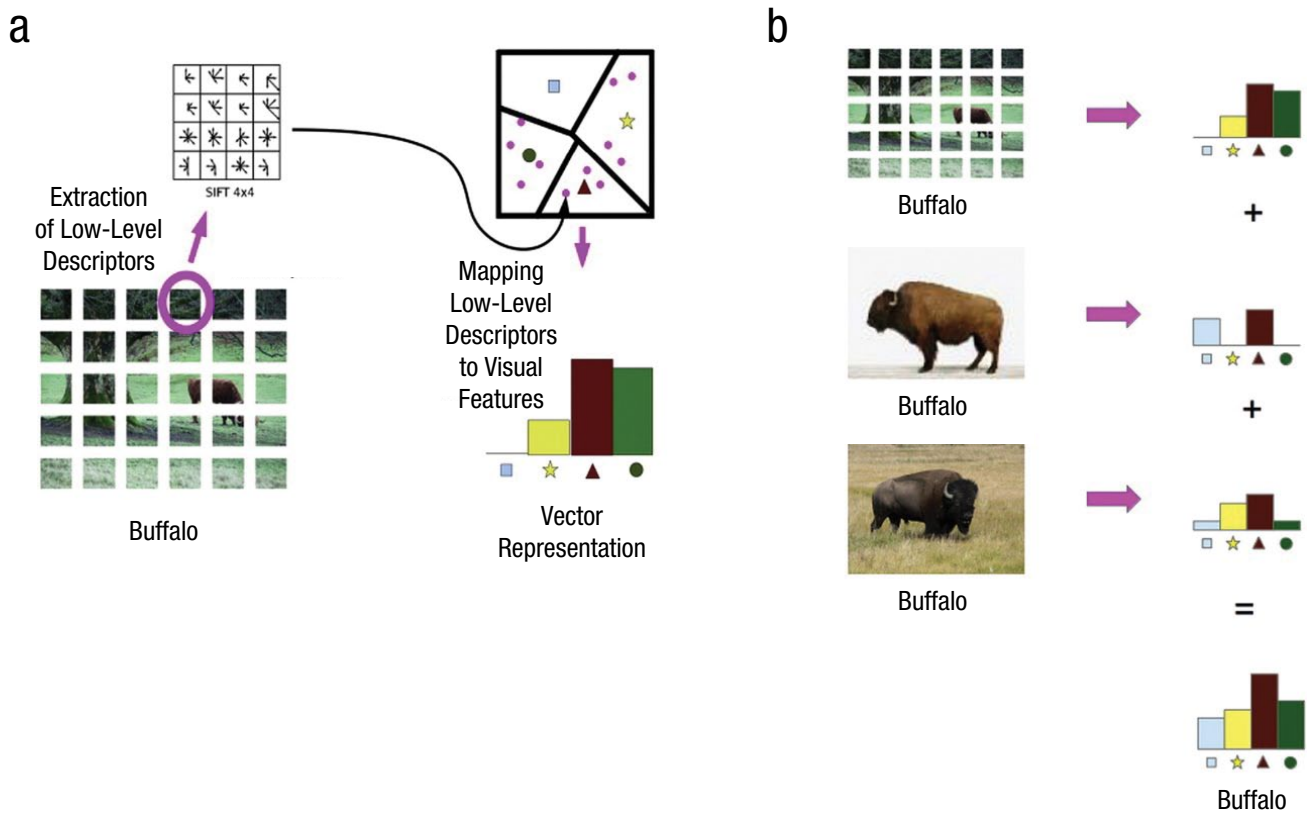
Fig. 2. Construction of a visual concept representation. When a single image is represented (a), low-level features of the image are extracted and then mapped to a set of higher-level visual features (which have been determined in advance by clustering low-level descriptors from a larger image collection). The image is represented by a vector that records (a function of) how often each visual feature occurs in it. The visual representation of a concept is shown in (b). Given a set of images depicting the same concept (e.g., a buffalo), the concept representation is obtained by summing the vectors representing all of the input images. SIFT = scale-invariant feature transform. Figure taken from Anderson, Bruni, Lopopolo, Poesio, and Baroni (2015).
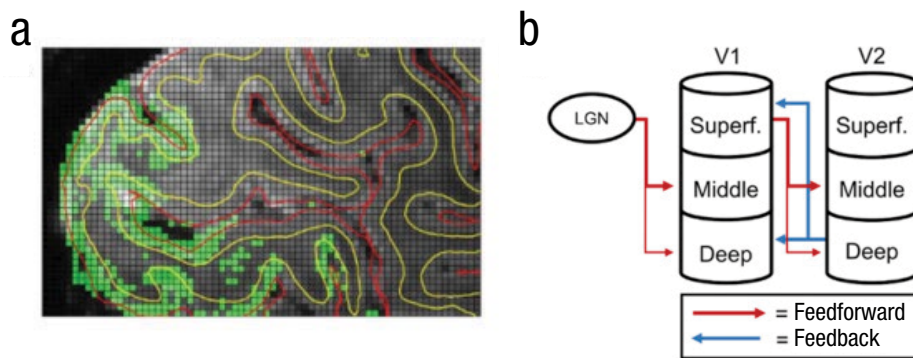
Fig. 3. Lamina-resolved functional MRI. In the sagittal slice of a functional volume (a), the black grid shows the size and location of 0.8-mm isotropic functional voxels. Boundaries of gray and white matter (yellow lines) and gray matter and cerebrospinal fluid (red lines) are overlaid onto the volume to show the distribution of functional voxels across cortical depths. The schematic (b) represents the feedforward connections from human lateral geniculate nucleus (LGN) to V1 and V2, as well as the feedback connections from V2 to V1. V1 and V2 are split into superficial, middle, and deep gray-matter layers to demonstrate how laminar functional MRI can be used to estimate feedforward and feedback responses by measuring layer-specific responses. Images and caption were adapted from (a) Kok, Bains, van Mourik, Norris, and de Lange (2016) and (b) Lawrence, Formisano, Muckli, and de Lange (2017).

Recommended Reading

Anderson, A. J., Bruni, E., Lopopolo, A., Poesio, M., & Baroni, M. (2015). (See References). A study demonstrating that activity patterns in high-level visual brain regions elicited by words can be predicted by image-based models that extract visual information from natural image statistics.

Barsalou, L. W. (1999). (See References). A classic article spelling out many of the key issues and questions that are still central aspects of current empirical and theoretical efforts today.

Mahon, B. Z., & Caramazza, A. (2008). (See References). A great reminder that alternative explanations (other than embodiment) can account for much of the empirical evidence and should thus not be ruled out too quickly.

Meteyard, L., Cuadrado, S. R., Bahrami, B., & Vigliocco, G. (2012). (See References). A great review that stresses which perspectives on embodiment are most likely to succeed.

Zwaan, R. A. (2014). (See References). An opinion article proposing a pluralistic view of conceptual processing in which different types of representations (such as embodied and amodal) engage in a division of labor.

Action Editor

Randall W. Engle served as action editor for this article.

ORCID iD

Markus Ostarek (iD) https://orcid.org/0000-0002-1138-7904

Acknowledgments

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

References

Anderson, A. J., Bruni, E., Lopopolo, A., Poesio, M., & Baroni, M. (2015). Reading visually embodied meaning from the brain: Visually grounded computational models decode visual-object mental imagery induced by written text. *NeuroImage*, *120*, 309–322.

Barsalou, L. W. (1999). Perceptions of perceptual symbols. *Behavioral & Brain Sciences*, *22*, 637–660.

Barsalou, L. W., Dutriaux, L., & Scheepers, C. (2018). Moving beyond the distinction between concrete and abstract concepts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1752), Article 20170144. doi:10.1098/rstb.2017.0144

Binder, J. R. (2016). In defense of abstract conceptual representations. *Psychonomic Bulletin & Review*, *23*, 1096–1108.

Borghesani, V., Pedregosa, F., Buiatti, M., Amadon, A., Eger, E., & Piazza, M. (2016). Word meaning in the ventral visual path: A perceptual to conceptual gradient of semantic coding. *NeuroImage*, *143*, 128–140.

Borghi, A. M., Binkofski, F., Castelfranchi, C., Cimatti, F., Scorolli, C., & Tummolini, L. (2017). The challenge of abstract concepts. *Psychological Bulletin*, *143*, 263–292.

Connell, L. (2007). Representing object colour in language comprehension. *Cognition*, *102*, 476–485.

Dove, G. (2016). Three symbol ungrounding problems: Abstract concepts and the future of embodied cognition. *Psychonomic Bulletin & Review*, *23*, 1109–1121. doi:10.3758/s13423-015-0825-4

Edmiston, P., & Lupyan, G. (2017). Visual interference disrupts visual knowledge. *Journal of Memory and Language*, *92*, 281–292.

Estes, Z., & Barsalou, L. W. (2018). A comprehensive meta-analysis of spatial interference from linguistic cues: Beyond Petrova et al. (2018). *Psychological Science*, *29*, 1558–1564.

Estes, Z., Verges, M., & Adelman, J. S. (2015). Words, objects, and locations: Perceptual matching explains spatial interference and facilitation. *Journal of Memory and Language*, *84*, 167–189.

Fernandino, L., Binder, J. R., Desai, R. H., Pendl, S. L., Humphries, C. J., Gross, W. L., . . . Seidenberg, M. S. (2016). Concept representation reflects multimodal abstraction: A framework for embodied semantics. *Cerebral Cortex*, *26*, 2018–2034. doi:10.1093/cercor/bhv020

Gozli, D. G., Chasteen, A. L., & Pratt, J. (2013). The cost and benefit of implicit spatial cues for visual attention. *Journal of Experimental Psychology: General*, *142*, 1028–1046.

Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, *458*, 632–635.

Kok, P., Bains, L. J., van Mourik, T., Norris, D. G., & de Lange, F. P. (2016). Selective activation of the deep layers of the human primary visual cortex by top-down feedback. *Current Biology*, *26*, 371–376.

Kouider, S., & Dehaene, S. (2007). Levels of processing during non-conscious perception: A critical review of visual masking. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*, 857–875.

Kousta, S.-T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: Why emotion matters. *Journal of Experimental Psychology: General*, *140*, 14–34.

Lakoff, G., & Johnson, M. (2003). *Metaphors we live by*. Chicago, IL: University of Chicago Press.

Lawrence, S. J., Formisano, E., Muckli, L., & de Lange, F. P. (2017). Laminar fMRI: Applications for cognitive neuroscience. *NeuroImage*, *197*, 785–791.

Lawrence, S. J. D., van Mourik, T., Kok, P., Koopmans, P. J., Norris, D. G., de Lange, F. P. (2018). *Laminar organization of working memory signals in human visual cortex*. Retrieved from SSRN: https://papers.ssrn.com/abstract=3155909

Lebois, L. A., Wilson-Mendenhall, C. D., & Barsalou, L. W. (2015). Are automatic conceptual cores the gold standard of semantic processing? The context-dependence of spatial meaning in grounded congruency effects. *Cognitive Science, 39*, 1764–1801.

Lupyan, G., & Ward, E. J. (2013). Language can boost otherwise unseen objects into visual awareness. *Proceedings of the National Academy of Sciences, USA, 110*, 14196–14201.

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.

Mannaert, L. N. H., Dijkstra, K., & Zwaan, R. A. (2017). Is color an integral part of a rich mental simulation? *Memory & Cognition, 45*, 974–982.

Merikle, P. M., Joordens, S., & Stolz, J. A. (1995). Measuring the relative magnitude of unconscious influences. *Consciousness and Cognition, 4*, 422–439.

Meteyard, L., Cuadrado, S. R., Bahrami, B., & Vigliocco, G. (2012). Coming of age: A review of embodiment and the neuroscience of semantics. *Cortex, 48*, 788–804.

Ostarek, M., & Huettig, F. (2017a). Spoken words can make the invisible visible—Testing the involvement of low-level visual representations in spoken word processing. *Journal of Experimental Psychology: Human Perception and Performance, 43*, 499–508.

Ostarek, M., & Huettig, F. (2017b). A task-dependent causal role for low-level visual processes in spoken word comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*, 1215–1224. doi:10.1037/xlm0000375

Ostarek, M., Ishag, A., Joosen, D., & Huettig, F. (2018). Saccade trajectories reveal dynamic interactions of semantic and spatial information during the processing of implicitly spatial words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*, 1658–1670.

Ostarek, M., Joosen, D., Ishag, A., De Nijs, M., & Huettig, F. (2019). Are visual processes causally involved in "perceptual simulation" effects in the sentence-picture verification task? *Cognition, 182*, 84–94.

Ostarek, M., & Vigliocco, G. (2017). Reading sky and seeing a cloud: On the relevance of events for perceptual simulation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*, 579–590.

Popper, K. (2005). *The logic of scientific discovery*. Abingdon, England: Routledge.

Pulvermüller, F., Hauk, O., Nikulin, V. V., & Ilmoniemi, R. J. (2005). Functional links between motor and language systems. *European Journal of Neuroscience, 21*, 793–797.

Tomasino, B., Fink, G. R., Sparing, R., Dafotakis, M., & Weiss, P. H. (2008). Action verbs and the primary motor cortex: A comparative TMS study of silent reading, frequency judgments, and motor imagery. *Neuropsychologia, 46*, 1915–1926.

Vukovic, N., Feurra, M., Shpektor, A., Myachykov, A., & Shtyrov, Y. (2017). Primary motor cortex functionally contributes to language comprehension: An online rTMS study. *Neuropsychologia, 96*, 222–229.

Willems, R. M., Labruna, L., D'Esposito, M., Ivry, R., & Casasanto, D. (2011). A functional role for the motor system in language understanding: Evidence from theta-burst transcranial magnetic stimulation. *Psychological Science, 22*, 849–854.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9*, 625–636.

Yang, Y., Dickey, M. W., Fiez, J., Murphy, B., Mitchell, T., Collinger, J., . . . Wang, W. (2017). Sensorimotor experience and verb-category mapping in human sensory, motor and parietal neurons. *Cortex, 92*, 304–319.

Zwaan, R. A. (2014). Embodiment and language comprehension: Reframing the discussion. *Trends in Cognitive Sciences, 18*, 229–234.

Zwaan, R. A., Etz, A., Lucas, R. E., & Donnellan, M. B. (2018). Making replication mainstream. *Behavioral & Brain Sciences, 41*, Article e120. doi:10.1017/S0140525X17001972

Zwaan, R. A., & Pecher, D. (2012). Revisiting mental simulation in language comprehension: Six replication attempts. *PLOS ONE, 7*(12), Article e51382. doi:10.1371/journal.pone.0051382

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