NeuroImage 99 (2014) 149–157

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Empirically grounding grounded cognition: The case of color

Ben D. Amsel [a,*], Thomas P. Urbach [a], Marta Kutas [a,b]

[a] Department of Cognitive Science, University of CA, San Diego, USA
[b] Department of Neurosciences, University of CA, San Diego, USA

ARTICLE INFO

Article history:
Accepted 10 May 2014
Available online 17 May 2014

ABSTRACT

Grounded cognition theories hold that the neural states involved in experiencing objects play a direct functional role in representing and accessing object knowledge from memory. However, extant data marshaled to support this view are also consistent with an opposing view that perceptuo-motor activations occur only following access to knowledge from amodal memory systems. We provide novel discriminating evidence for the functional involvement of visuo-perceptual states specifically in accessing knowledge about an object's color. We recorded event-related brain potentials (ERPs) while manipulating the visual contrast of monochromatic words ("lime") in a semantic decision task: responses were made for valid colors ("green") and locations ("kitchen") and withheld for invalid colors and locations. Low contrast delayed perceptual processing for both color and location. Critically, low contrast slowed access to color knowledge only. This finding reveals that the visual system plays a functional role in accessing object knowledge and uniquely supports grounded views of cognition.

Introduction

The long-standing controversy about how knowledge is retained and accessed in the human brain remains unresolved. On the amodal views that emerged alongside the artificial intelligence and cognitive revolutions of the 20th century, knowledge representations (i.e., concepts) are modality independent, no longer traceable to the neural states leading to their initial perceptual and/or motor encoding (Caramazza et al., 1990; Landauer and Dumais, 1997; Levelt, 1989; Pylyshyn, 1984; Riddoch et al., 1988). In other words, all that one may know about a thing (e.g., its weight, odor, function, location, and color) is represented via a neural code irrevocably de-coupled from the perceptuo-motor encoding systems that serve actually "experiencing" that thing. Consequently, even if different kinds of information about a thing may be more or less difficult to access (e.g., due to competition for limited attentional and/or working memory resources), knowing the weight, odor, or color of an object is not mediated by the somato-motor, olfactory, or visual systems, respectively. In short, on amodal views, a person can access their knowledge that limes are green without engaging their visual perceptual system, and conceptual decisions regarding perceptual object attributes should not be directly affected by variations in those perceptual attributes.

Alternatively, on grounded views they should. Grounded views of cognition hold that at least some of the neural resources responsible for perceiving and acting upon an object play a functional role in representing and accessing knowledge of that object in long-term memory (Allport, 1985; Barsalou, 1999, 2008; Martin, 2007; Paivio, 1986; Pulvermüller, 2013; Warrington and Shallice, 1984). On grounded views, perceptual (and motor) systems are not merely the conduits of sensory (and motor) experiences to (and from) amodal semantic memory systems, but rather play a functional role in the storage of and access to perceptuo-motor knowledge about objects. From this it follows that access to certain kinds of knowledge about an object will be differentially subject to the specific properties and sensitivities of modality-specific neural resources involved, even when access is prompted by a word (referring to the object) and not by perception of the object itself. In short, on grounded views, a person would not explicitly access their knowledge that limes are green, for example, without engaging their color processing system, and the nature of that semantic access (e.g., timing, realization) would be selectively impacted by any perceptual variable that modulates that particular perceptual system.

Amodal and grounded views comprise fundamentally different accounts of the functional organization of the human brain, and give rise to distinctly different experimental predictions. And, yet, an impasse between these theoretical positions persists because upon closer inspection data marshaled as evidence for or against each is consistent with one view or the other but is not discriminative between them.

Both grounded and amodal views are reasonable based on available data. Proponents of amodal views, for example, point to the activation of modality-independent (rather than modality-specific) neural systems during verb comprehension (Bedny and Caramazza, 2011), and to the observation that sighted versus congenitally blind individuals exhibit statistically indistinguishable functional and spatial profiles of fMRI activation in putative high-level visual brain regions during size or similarity judgments about the referents of spoken words (Bedny et al., 2012; Mahon et al., 2009). Given that blind individuals do not have direct visual experiences with objects, these results would seem to be

---

* Corresponding author.
E-mail address: bamsel@ucsd.edu (B.D. Amsel).

http://dx.doi.org/10.1016/j.neuroimage.2014.05.025
1053-8119/

strong evidence for an amodal view. However, the information needed to judge the size or weight of an object, for example, can be acquired via tactile experience and not just visual experience, and the occipitotemporal areas of blind individuals can undergo substantial compensatory reorganization such that these ventral stream "visual regions" (at least in sighted individuals) could be recruited for somatosensation in blind individuals.

In the case of color, lesions to ventral temporal cortex (typically fusiform gyrus) are associated with intact color perception despite impaired access to color knowledge (Miceli et al., 2001) whereas lesions to lingual gyrus can lead to the opposite pattern of dissociation (i.e., achromatopsia: Bouvier and Engel, 2006). These results rule out an identical neural mechanism for perceiving color and accessing color knowledge. Crucially, however, they do not rule out the possibility of at least some common mechanisms between color perception and color knowledge and thus do not constitute discriminative evidence for the amodal view.

Proponents of grounded views point to several behavioral, imaging, and lesion studies to make their case. Behavioral modality-specific interference effects have been taken as support for a functional link between perceptual processes and semantic memory access (Boulenger et al., 2006; Kaschak et al., 2005; Meteyard et al., 2007, 2008; Pecher et al., 2003; van Dantzig et al., 2008). For example, verification of modality-specific object properties (e.g., "a bee buzzes") is slower following a conceptual judgment about a property in a different modality (e.g., "a lime is green") (Pecher et al., 2003), and following a perceptual judgment in a different modality (e.g., whether an LED was illuminated in the left or right visual field) (van Dantzig et al., 2008). Perceived motion in a specific direction (e.g., dots moving upwards) slowed lexical decision latencies on motion words (e.g., "fall") that were incongruent with the perceived motion direction (Meteyard et al., 2008). Conversely, visual motion perception can be impacted by motion-related language processing (Dils and Boroditsky, 2010; Meteyard et al., 2007). These behavioral effects are consistent with grounded views, but they also could reflect interactions between perceptual and language-related processes subsequent to semantic access, and therefore do not definitively discriminate between amodal and grounded views of memory access and representation (Hauk and Tschentscher, 2013; Mahon and Caramazza, 2008).

Transcranial magnetic stimulation of the leg area of somatotopically-organized motor cortex led to decreased lexical decision times for leg-related words (e.g., "kick") but not arm-related words (Pulvermuller et al., 2005); these findings were interpreted as demonstrating a specific functional link between action and lexico-conceptual systems. A similar study (Tomasino et al., 2008), however, replicated this finding only when participants imagined performing the actions but not when the verbs were silently read—findings which were taken as evidence against grounded views. Proponents of grounded views also point to functional magnetic resonance imaging (fMRI) activations of visual, olfactory, and gustatory brain regions during perceptual tasks that require access to knowledge about these sensory properties from written words (Barros-Loscertales et al., 2012; Chao and Martin, 1999; Gonzalez et al., 2006; Hsu et al., 2011; Simmons et al., 2007). In these studies, activated voxels during judgments about the colors of objects from their names alone could be proximal to, or overlapping with, activated voxels during color perception (Chao and Martin, 1999; Hsu et al., 2011; Simmons et al., 2007). Although these findings are compatible with grounded views, it has been persuasively argued that they do not conclusively establish its basic premise that these activated perceptual systems play a *functional role* in accessing conceptual knowledge (Mahon and Caramazza, 2008; Rugg and Thompson-Schill, 2013). The relatively poor temporal resolution of fMRI makes it impossible to determine whether perceptual regions activated during conceptual processing reflect perceptual mediation of semantic access or top-down feedback from amodal representations subsequent to semantic access (Chan et al., 2011; Halgren et al., 2006).

Electrophysiological results consistent with early engagement (<250 ms following stimulus onset) of sensory or motor systems during language comprehension (Amsel, 2011; Boulenger et al., 2006; Hoenig et al., 2008; Moscoso del Prado et al., 2006; Pulvermüller et al., 2009) do not settle the matter either. Although these measures have the requisite temporal resolution, they do not allow strong inferences as to the nature of the processes—perceptual and/or cognitive—interposed between the perception of a word form and the associated early activation of perceptual or motor systems (Mahon and Caramazza, 2008).

At present, the debate between amodal and grounded views is at an impasse. Continued research along now familiar lines will likely be consistent with one view or another, but is unlikely to lead to clear resolution of the fundamental question (Chatterjee, 2010; Mahon and Caramazza, 2008). Our approach to this impasse was to consider where amodal and grounded views make fundamentally different predictions, and what kind of empirical results could (possibly) constitute definitive evidence for either set of predictions. In our view, a demonstration that a perceptual neural processor (e.g., visual or auditory system) plays a functional role in access to specific object knowledge (e.g., color or sound knowledge) would be incontrovertible evidence for grounded views. The key objective would be to demonstrate that a specific perceptual system is not only engaged relatively early in a conceptual task reliant on access to that perceptual knowledge, but is specifically recruited in the service of semantic access of that knowledge. We will show that these desiderata are met by an electrical brain measure—event-related brain potentials (ERPs) in the context of a novel experimental design that meets this challenge.

We assessed the extent to which a purely visual manipulation would selectively impact access to knowledge about an object's typical color from its written name. We varied the relative luminance of monochromatic words against a constant background (i.e., visual contrast) to modulate activity in visual systems sensitive to luminance and color. Increasing visual contrast of words and images modulates neural activity in several hierarchical visual regions (Avidan et al., 2002; Kastner et al., 1999; Mechelli et al., 2000; Tootell et al., 1995) including ventral stream regions (e.g., fusiform gyrus) implicated in color processing (Grill-Spector and Malach, 2004). At issue is whether varying the visual contrast of written words would differentially affect the timing of access to color knowledge (the referent's typical color) versus location knowledge (the referent's typical location).

Unlike images depicting objects, understanding written words requires access to knowledge from an arbitrary symbol (i.e., the word form). The physical properties of these symbols bear no relation to the physical properties of their referents (the monochromatic printed word "lime" has no relation to a lime's greenness). Whereas color-selective fMRI activity triggered by a word (e.g., an object name) is silent as to whether the voxel was activated during or subsequent to access of color knowledge, a demonstration that the timing of access to color knowledge is modulated by a purely visual variable (visual contrast) would constitute evidence uniquely for the grounded view. Alternatively, whether a written word (e.g., "lime") appears in high or low visual contrast should have no impact on the time it takes to verify the word referent's typical color (e.g., green, purple) or location (e.g., kitchen, pond), or perforce, if it does, then according to amodal views, it should not do so differentially. In sum, although (pre-conceptual) sensory processing may be delayed for low relative to high contrast words, according to architectures in which color and location knowledge reside in amodal memory systems that are (by definition) impervious to changes in visual contrast, our purely visual manipulation should have either no effect or no differential effect on the timing of access to different kinds of knowledge.

To test these different predictions we chose an electrical brain measure that indexes various perceptual, conceptual, and motor processes as they unfold in time — the average scalprecorded event-related brain potential (ERP). The ERP waveform reflects instantaneous summed post-synaptic potential activity primarily of pyramidal cells in neocortex time-locked to an eliciting event; moreover, different segments of the ERP waveform reflect qualitatively different neurocognitive processes. Potentials within the first ~200 ms, such as the P2 component, are sensitive to physical stimulus parameters and thus will be delayed to words viewed under low (versus higher) visual contrast—regardless of the task or the words' meanings. To determine whether or not (and if so the extent to which) the timing of semantic access also is affected by low visual contrast, we employed a go/nogo semantic decision task.

We recorded ERPs to word pairs (object property–object name) as participants performed a go/nogo semantic decision task in which responses were contingent upon whether a color (e.g., “green” or “purple”) or location (e.g., “kitchen” or “pond”) is a valid or invalid attribute of an object (e.g., “lime”). In comparison with ERPs elicited when responding to a valid attribute (go trials), withholding a response (nogo trials) elicits a more negative ERP component (nogo N200) originating in frontal cortex (Lavric et al., 2004; Sasaki et al., 1993). The onset latency of the derived N200 effect (i.e., the time by which the nogo ERP reliably diverges from the go ERP) is taken as an upper limit of when sufficient information has become available from a stimulus to determine whether or not to make a response. Specifying the exact nature of the neural mechanisms of the go and nogo ERPs is not necessary for this inference, nor is specifying which of the two waveforms is changing relative to the other across experimental conditions. This paradigm has been used to pinpoint the timing of different processes in language production (Schmitt et al., 2000, 2001), and to infer how quickly the brain can distinguish between images (Thorpe et al., 1996; VanRullen and Thorpe, 2001) or written words (Amsel et al., 2013; Hauk et al., 2012) denoting living versus nonliving things.

With the latency of N200 effect indexing the upper bound of semantic access, the amodal view predicts that the N200 effect should be the same (Fig. 1A), or equally delayed for both types of decisions (Fig. 1B). On the other hand, the grounded view maintains that access to knowledge about a referent's color directly involves regions of visual cortex that remain sensitive to visual contrast, whereas access to knowledge about an object's typical location involves these regions less so or not at all (Fig. 1C). As a consequence, if grounded views are correct then words presented under low versus normal contrast should affect the timing of access to color knowledge more than the timing of access to location knowledge. Thus, to the extent that visual contrast differentially impacts the timing of access to color versus location knowledge about the referent of a written word, the onus will be on opponents of grounded views to explain how this would be possible in a strictly amodal architecture.

Method

Participants

Twenty-five right-handed (Oldfield, 1971) undergraduate students (15 males) were recruited from the University of California, San Diego. Participants were native English speakers between 18 and 26 years of age (M = 20.5, SD = 2.1) with no exposure to other languages before age seven. Participants had normal or corrected-to-normal vision, and reported no major neurological or general health problems or psychoactive medication use. Each participant provided written informed consent prior to the experiment and received course credit and/or $7/h for participating.

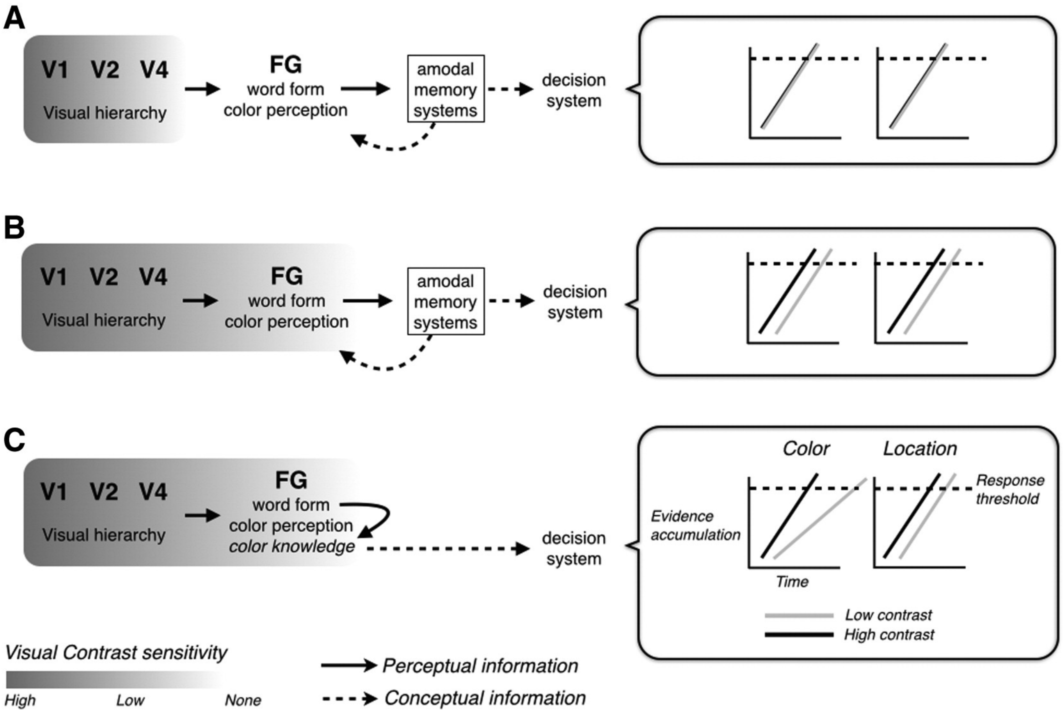

Fig. 1. Potential outcomes in amodal and grounded architectures. Each panel depicts an amodal (A, B) or grounded (C) architecture and the processes underlying a color verification decision. From left to right, contrast-dependent low-level visual regions propagate visual form information to the fusiform gyrus (FG), where visual word form processing becomes view-invariant. Some time after the word form is processed, conceptual information becomes available to a decision-making system that can signal the motor system to execute or withhold a response in the go/nogo task. The solid lines in the right-most panel schematically depict the rate of evidence accumulation for color and location decisions and the dotted line indicates the threshold for signaling the response. In this experiment the onset latency of the N200 ERP effect provides an upper bound on the time by which this response threshold is reached. A) An amodal architecture wherein word form information is propagated eventually to amodal memory systems, where semantic access provides evidence to a decision system that signals the response. B) Identical to panel A with the exception that FG remains sensitive to the visual contrast of written words, resulting in a constant delay in low (versus high) contrast color and location decisions. C) A grounded architecture wherein word form information in FG is directly routed to nearby or overlapping FG tissue involved in color perception and also access to color knowledge. Low contrast text causes a sensory-based equivalent slowing of access to color and to location knowledge, but access to color knowledge is additionally slowed by the direct effect of low contrast on the neural resources involved in accessing color knowledge. Note that although it is not depicted here, we assume that feedback may occur at some or all levels of these architectures.

Stimuli

One hundred and eighty objects were paired with one valid color and location, resulting in 360 property–concept pairs. Thirteen color names ("black", "blue", "brown", "gold", "gray", "green", "orange", "pink", "purple", "red", "silver", "white", "yellow") were selected from a large set of feature production norms (McRae et al., 2005) wherein each object name was presented and at least 7 of 30 participants produced the color name used in our study (e.g., "is red"). Words denoting twenty-six locations ("backyard", "bathroom", "battlefield", "concert", "desk", "dresser", "farm", "forest", "fridge", "garage", "garden", "grass", "house", "jungle", "kitchen", "ocean", "party", "pond", "prairie", "street", "swamp", "tree", "tropics", "water", "winery", "zoo") were selected in part from the same dataset and in part by the experimenters. Each color name and location name was paired with between 2 and 33 ($M = 10.5$, $SD = 8.4$) object names. We created an additional 360 invalid concept–property pairs (e.g., blue lime) by shuffling the valid pairs. We distributed the invalid properties as much as possible across different objects to avoid any association between valid and invalid pairs (e.g., concepts paired with the color name "green" were paired with a variety of different location properties in the invalid pairs). We attempted to minimize the difficulty of determining whether a trial was valid or invalid by avoiding semantically similar pairs (e.g., if valid pairs were "yellow banana" or "kitchen banana" the invalid pairs would not be "orange banana" or "bedroom banana").

The stimulus words were displayed on a CRT monitor (ViewSonic P220f). Visual stimulus contrast was manipulated by presenting the text in one of two shades of lighter gray against a darker gray background. Luminance, hue, and saturation (CIE xyY color space) were measured with a Konica Minolta Chroma Meter (CS-100A) at the location of stimulus presentation. The luminances of the slightly lighter low contrast text (47.6 cd/m^2) and the substantially lighter high contrast text (57.3 cd/m^2) were approximately 5% and 25% higher, respectively, than the luminance of the background (45.4 cd/m^2). The hue (range = .298–.299) and saturation (.314) were essentially identical for the background and both foreground grays. Measurements made following each testing session revealed that luminance varied ± 0.3 cd/m^2 (i.e., <1.5%) and there was no measureable variation in hue or saturation. Stimulus words were presented slightly above center in a Helvetica font where each character subtended about 0.8° of visual angle in height and 0.6 in width.

Design

We created a 2 (response: go, nogo) × 2 (decision: color, location) × 2 (visual contrast: high, low) factorial within-subjects design. Each participant performed four blocks in which they responded (go trials) either to valid pairs or invalid pairs, and made decisions about either colors or locations. Decision type (color/location) was blocked, whereas contrast (high/low) was randomized within blocks. Each of the 180 concepts appeared once in every block. Two versions of the stimuli were created to balance the go/nogo criterion across all property–concept pairs. Participants were asked to respond (go trials) to invalid trials in half of the blocks to deter response strategies based on validity. Following the seminal go/nogo ERP study in this paradigm (Thorpe et al., 1996), behavioral and ERP data were analyzed from trials in which participants responded (go trials) to valid pairs and withheld responses (nogo trials) to invalid pairs.

Visual contrast was split evenly within and across blocks as well as within concepts. For example, if "green lime" was presented in low contrast when the response criterion was to 'go' to valid trials, it would be presented in high contrast when the criterion was to 'go' to invalid trials. This relationship was reversed for half of the trials to ensure that all manipulated variables were in fact independent. Within each of the two lists, a unique permutation of block order was presented to each participant. Within each block the order of trial presentation was selected at random with the exception that trials requiring a given response type (i.e., go or nogo) never appeared in succession more than four times.

Procedure

Participants were tested individually while seated in a dimly lit, sound attenuating, electrically-shielded chamber, in front of the CRT monitor at a viewing distance of 112 cm. At the beginning of the experiment the participant was shown seven word pairs in low contrast, and asked to name each of them aloud to ensure that the low contrast trials were visible. Before each experimental block the experimenter explained the decision criterion, showed the participant some examples of valid and invalid trials in high and low contrast, and ensured that the participant understood the correct decision for each. Finally, the participant completed 26 practice trials identical to the experimental trials with the exception that the experimental and practice concepts did not overlap. Thirteen trials were shown in low contrast and 13 in high contrast, and every experimental property appeared at least once.

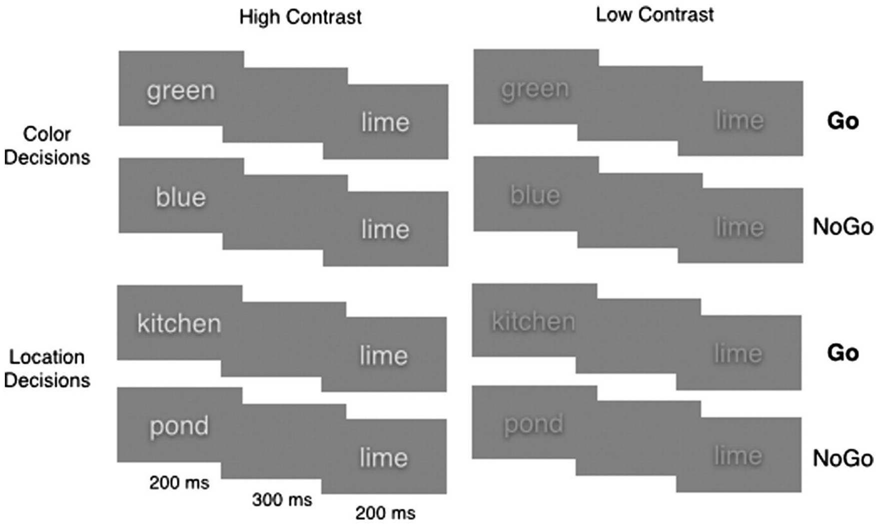

Fig. 2. Depiction of stimuli and trial timing. High and low contrast stimuli shown here are not veridical depictions of actual experimental stimuli; please see Methods for exact stimulus parameters (i.e., luminance, hue, saturation, size, etc.).

The experimental trials in each block were separated by rest periods, indicated by a message that remained on the screen until the participant elected to continue. Each block lasted approximately 15 min and the entire experiment typically lasted less than 2.5 h. Each trial began with the first word (property) for 200 ms, followed by a 300 ms blank screen, followed by the second word (concept) for 200 ms. A blank screen then appeared for a randomly selected interval between 2200 and 2400 ms (See Fig. 2). The words appeared above a small gray fixation square subtending about 0.5° of visual angle in height and width, that remained on the screen throughout each trial. Participants were instructed to rest their arms on their laps and rest their right thumb on a response button mounted on a rubber handle. Participants were asked to refrain from blinking and other movement between the onset of the first word and at least 1 s following the onset of the second word.

Apparatus and recording

Response latencies were measured from the onset of the object name, and responses occurring after 2000 ms were not registered. The electroencephalogram (EEG) was continuously recorded from 26 geodesically-arranged tin electrodes (Ganis et al., 1996) embedded in an ElectroCap (impedances were kept below 5 kΩ), and referenced to the left mastoid. Eye movements and blinks were monitored with electrodes placed on the left and right lower orbital ridges, and left and right external canthi. The EEG was digitized at a sampling rate of 250 Hz and bandpass filtered between 0.01 and 100 Hz with James Long amplifiers (www.JamesLong.net). Potentials were re-referenced offline to the mean of the left and right mastoid electrodes. Averages were obtained for 1000 ms epochs including a 200 ms baseline period prior to stimulus onset. Trials of correct responses were visually inspected for each subject. Trials containing eye movements, amplifier blocking, or any other artifacts within the critical time window were discarded using individualized rejection criteria. Overall, 17.2% of all trials were discarded due to artifacts. The number of rejections did not reliably differ between the decision categories or between high and low contrast trials. The percentage of rejected go trials (11.3%) was significantly lower than the percentage of rejected nogo trials (23.2%), $F(1, 24) = 24.9$, $p < .01$.

Statistical analysis of ERPs

We quantified the onset latency of the N200 effect by submitting the amplitude of the nogo and go ERP in each decision condition to repeated measures t-tests at every time point between 100 and 600 ms at 11 frontal and prefrontal scalp sites that encompass the nogo N200 component (1375 comparisons per condition). Protection against a large proportion of false discoveries without excessive loss of statistical power was provided by the adaptive two-stage false discovery rate control (FDR) procedure (Benjamini et al., 2006), thereby determining which t-tests were significant at an FDR level of 0.05. Simulation studies have showed that this procedure provides accurate control of the FDR rate for ERP data (Groppe et al., 2011).

Results

Behavior

Property verification accuracy

Not surprisingly, accuracy was overall higher on high versus low visual contrast trials for both valid go trials (87.5% vs. 83.0%, respectively, $F(1, 24) = 13.3$, $p < .01$) and invalid nogo trials (89.8% vs. 84.0%, $F(1, 24) = 18.3$, $p < .01$). Accuracy did not differ reliably for color versus location decisions for the valid go or invalid nogo trials, nor did this factor interact with visual contrast on either type of trial. Sensitivity (d') was significantly higher for high contrast (3.1) versus low contrast (2.6) trials, $F(1, 24) = 27.6$, $p < .01$, but did not differ across color versus location decisions, nor was there a significant interaction of the nature of the decision with visual contrast.

Correct go response times

All correct go responses greater than 200 ms and less than 2000 ms were retained for analysis. High contrast trials ($M = 843$ ms, $SD = 142$ ms) were 55 ms faster than low contrast trials ($M = 898$ ms, $SD = 155$ ms), $F(1, 24) = 33.7$, $p < 0.01$. There was no significant main effect of the type of property verification decision, i.e., response times for correct color ($M = 876$ ms, $SD = 154$ ms) and location ($M = 865$ ms, $SD = 148$ ms) verifications were not reliably different, $F(1, 24) = 0.7$, $p > .4$. The interaction effect between decision type and visual contrast did not reach significance, $F(1, 24) = 3.3$, $p = .08$; under high contrast the color and location trials were nearly equal (2 ms difference, $t(24) = -.14$, $p = .90$), whereas under low contrast response latencies for color trials were 26 ms longer than for location trials, $t(24) = 1.9$, $p = .07$.

Event-related brain potentials (ERPs)

Time domain average ERPs relative to mean amplitude 200 ms prestimulus were computed for each participant, separately for high and low contrast go trials (valid properties) and high and low contrast nogo trials (invalid properties). The early sensory N1 and P2 components are part of the obligatory neural response to visual stimuli, and are delayed by about 50 ms in the low versus high contrast conditions for the first words and second words for both decision types (Fig. 3).

Fig. 4 depicts relevant go and nogo ERPs for color and location decisions. The onsets and slopes of the initial nogo minus go N200 difference ERPs for low contrast versus high contrast trials are very similar in the location decision task (i.e., between about 200 and 500 ms). By comparison, in the color trials the low contrast difference wave reliably diverges from zero substantially later than the high contrast difference wave; its slope is visibly less steep. We quantified these latency differences using mass univariate analyses to determine the earliest time point by which the nogo ERPs and go ERPs diverge from each other (i.e., when the amplitude of the N200 difference ERP in each decision condition was reliably larger than zero). This measure was computed separately for each condition (Fig. 5) and we operationally defined the onset of the N200 effect as the first time point by which a statistically significant difference was obtained at a minimum of 15 previous consecutive time points (60 ms), for at least two pre/frontal electrode sites. On this

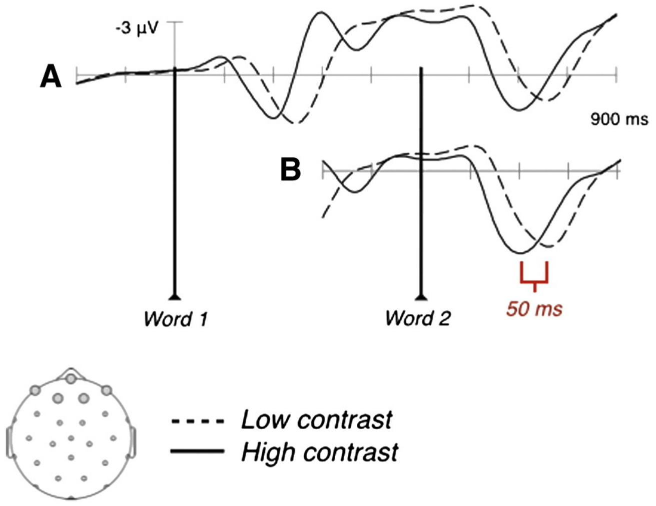

Fig. 3. Visuo-perceptual ERP components are not influenced by semantic decision type. The P2 ERP component is an obligatory neural response to visual stimuli. Low contrast delays the P2 component by about 50 ms for the first word (A) and second word (B) of the word pairs. Importantly, the 50 ms delay induced by low contrast (text) is not significantly different for color and location decisions.

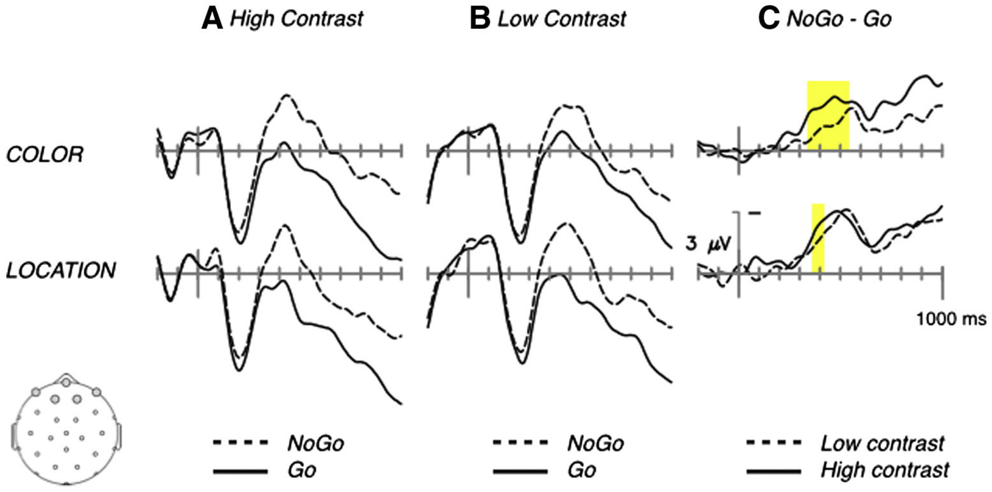

Fig. 4. Differences between go and nogo ERP waveforms reveal a differential delay in access to color knowledge. ERP waveforms averaged across five electrode sites over frontal cortex (legend in the bottom left corner) for high contrast trials (Panel A), low contrast trials (Panel B), and for nogo–go difference ERPs (N200 effect; Panel C). Note the delay in the low versus high contrast N200 effect in the color decisions in comparison with the location decisions (yellow shading; panel C).

analysis, the difference between the onset of a reliable N200 effect in high contrast color verification trials (332 ms) and low contrast color verification trials (536 ms) is 204 ms, whereas the difference between the onsets of high contrast location trials (368 ms) and low contrast location trials (408 ms) is 40 ms. Alternate methods of computing these latencies also yielded a substantially later onset for low contrast color trials only.

Discussion

We asked people to make property verification judgments for pairs of written words presented under either normal or low visual contrast as a means of assessing the differential predictions of amodal versus grounded views of knowledge representation in the human brain. Half of the decisions were about the word referent's typical color and half about its typical location; according to amodal views the nature of the semantic decision should not interact with the visual contrast of the written words whereas on grounded views it should. More specifically, if the visual system plays a functional role in accessing color knowledge then this should be reflected in the timing of any neural process that relies on decisions that engage those perceptual processes; in the current experiment, this holds for decisions about an object's color even though the object is referred to via a monochrome written word, because the color decision is known to engage brain areas sensitive to visual contrast. We inferred the timing of semantic access and decision processes from the time course of scalp recorded brain activity—the N200 effect—a difference between the brain's response to the words on go trials (for valid object properties) and nogo trials (for invalid object properties).

The results were clear. Access to knowledge about an object's typical location was delayed by about 50 ms when the words were presented under low visual contrast, reflecting the 50 ms delay in the early visual potentials evoked by those words. This result is consistent with both amodal and grounded views. Access to object knowledge of the referent's typical color, however, was approximately 150 ms later when the word pairs were presented in low visual contrast. This additional timing difference persisted to a lesser degree in the peak latency of a later ERP component thought to reflect stimulus evaluation (P3b on go trials), and in the subsequent behavioral response times for the (valid) go trials (Fig. 6). The larger timing differences in the N200 effects compared to the reaction time effects are not without precedent in this paradigm (Rodriguez-Fornells et al., 2002; Schmitt et al., 2000). Our data indicate a substantial timing difference in semantic access for color verification in comparison to location verification above and beyond the perceptual delay seen in the sensory evoked potentials. This timing difference between the N200 effects does not appear to be a simple latency shift in the difference waveform, as is the case in the earlier P2 component. Rather, the N200 effects appear to be characterized by different slopes, which may reflect differences in the rate of accumulation of evidence for the semantic decisions. Critically, whatever neural mechanisms underlie these timing differences, this specific pattern of timing differences uniquely supports grounded views of knowledge representation.

Amodal memory systems are, by definition, insensitive to purely perceptual stimulus properties and thus should not be affected by visual contrast. Low versus normal visual contrast should not interact differentially with the time it takes to access different kinds of knowledge (e.g., location vs. color) from an amodal memory system. However, we show that it does and in a way that recourse to post-access, top-down effects on perceptual processes from an amodal memory system cannot readily explain (see Figs. 1A & B). Instead, our findings indicate a direct functional role of the color processing perceptual system in accessing color knowledge from memory—the very challenge posed against any grounded account and met by these data.

Importantly, we can rule out potential stimulus or task-related confounds. Our participants were equally accurate in their location and color decisions regardless of visual contrast. Location words were longer on average than color words, and the set size of the locations was twice that of the colors. If these stimulus differences influenced our results, they should have worked against our finding of longer nogo N200 latencies and response times for color relative to location decisions. The number of letters that can be decoded in a single fixation is known to decrease with decreasing visual contrast (Legge et al., 1997), and increasing set size typically slows decision latencies. Moreover, a post-access mental imagery account is further incompatible with the timing of the N200 effect which is visible by 300 ms for the high contrast color decisions—before the documented 500 to 800 ms time window in ERP studies of mental imagery generation (Gullick et al., 2013; West and Holcomb, 2000).

Our results are consistent with direct functional overlap (but not necessarily neuroanatomical overlap) between the processes of visual word form analysis, color perception, and conceptual access to color knowledge. Functional overlap among grapheme-sensitive, color-sensitive, and contrast-sensitive regions of the ventral temporal lobe is supported by experiments with grapheme–color synesthetes who experience monochromatic graphemes in specific colors, and for whom the strength of experienced color intensity monotonically decreases with decreasing visual contrast (Hubbard et al., 2006). Grapheme–color synesthetes exhibit greater structural connectivity and gray

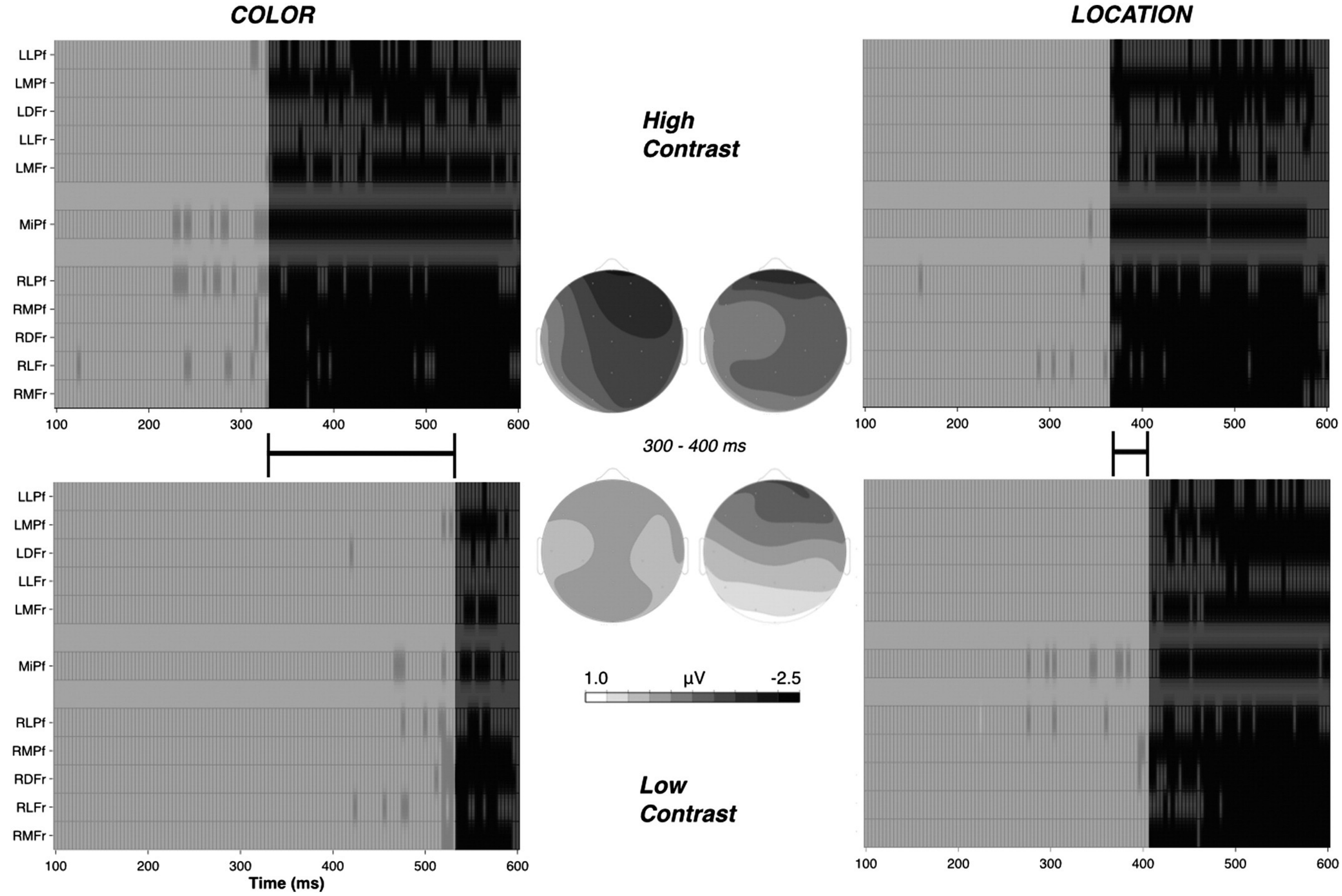

Fig. 5. The N200 ERP effect is substantially more delayed by low contrast text for color decisions versus location decisions. Raster plots (outer four plots) show the results of repeated measures t-tests computed at every time point between 100 and 600 ms at 11 frontal and prefrontal scalp sites. t-Tests that are statistically significant at an FDR level of 0.05 are shaded in black. Laterality (left, middle, right) is represented by the top, middle, and bottom sections of each raster plot. The onset of a reliable N200 effect is defined here as the first time point after which statistically significant t-tests occur for fifteen or more consecutive time points (60 ms window) at two or more electrode sites, and is depicted by the border between the shaded and non-shaded halves of each raster plot. Black "bar-bells" visible between the high and low contrast plots depict the difference in milliseconds between low and high contrast trials for the color decisions (left side; 204 ms) and location decisions (right side; 40 ms). Scalp maps (inner four plots) show spherical spline interpolated distributions of the grand average nogo–go ERP difference (N200 effect) between 300 and 400 ms. By this latency the frontally-distributed N200 effect is visible for both decision types under high contrast, but for location decisions only under low contrast.

matter volume in inferior temporal regions involved in visual word form and color perception (Rouw and Scholte, 2007; Weiss and Fink, 2009), consistent with hyper-connectivity between these visual regions during grapheme perception (Brang et al., 2010; Ramachandran and Hubbard, 2001). An exciting future prospect would be to examine grapheme–color synesthetes with the current design. Synesthetes could be expected to incur an additional delay in accessing color knowledge beyond the delay reported here, due to additional interference from synesthetically-experienced colors of monochromatic graphemes.

An important question raised by our results is why access to location knowledge did not incur any additional delay beyond the perceptual delay indexed by early visually-evoked waveforms. One possibility is that there is functional overlap between color perception and access to color knowledge, but not between other kinds of visual perceptual processes and the processes involved in accessing location knowledge. In this case, access to location knowledge could involve aspects of the visual system not sensitive to our contrast manipulation. Another possibility is that access to knowledge about an object's typical location exclusively involves an amodal memory system, in which case no perceptual manipulation could influence the time course of memory access beyond a perceptual delay. It is important to note that we are not arguing that all conceptual knowledge is represented and accessed in a grounded fashion; it remains to be seen what knowledge, and under what circumstances, fits this description.

Finally, we speculate that a neural mechanism for the present result could be temporal coding by response latency. At the single neuron level, response latency refers to the timing of spikes relative to stimulus onset. Studies of the macaque visual system have shown that decreasing visual contrast causes an increasing delay in neural response latency, and that the delay becomes progressively longer in higher-order (more anterior) visual areas (Oram et al., 2002). An intriguing possibility awaiting further study is that increasing perceptual difficulty (e.g., decreasing visual contrast) delays the firing onset of neurons that play a functional role *both* in visual perception and in access to visual knowledge.

In conclusion, we used visual contrast—a manipulation to which amodal memory stores are immune but visual perceptual processes are sensitive—to show that the visual processing system plays a direct functional role in accessing conceptual knowledge about a referent object's color from a monochromatic word. Our results unequivocally support grounded views of knowledge representation and undermine strict amodal views. It seems then that deciding that limes are green (i.e., the word "green" appropriately describes the word "lime") directly involves the visuo-perceptual system with which the greenness of limes is perceptually experienced. Is that perceptual activation a prerequisite for understanding the word "lime", even if that knowledge was gained secondhand, e.g., from a book? We cannot yet be sure, but moving forward it will be important to determine the boundary conditions of

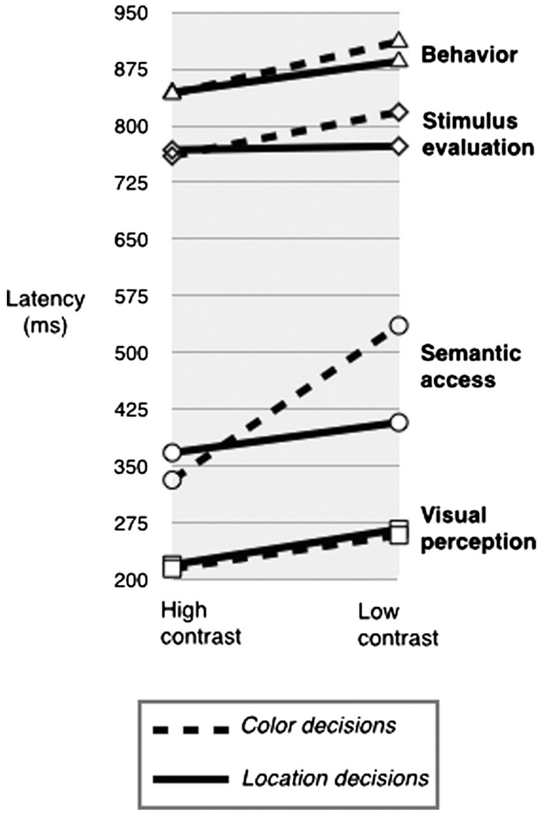

Fig. 6. Differentially delayed access to color knowledge persists from a relatively early marker of semantic to behavior. Visual perception is operationalized as the peak latency of the visual P2 component during go trials, and is delayed by a constant amount in all low contrast trials. An upper bound of semantic access is operationalized as the onset of a significant N200 ERP effect (see Fig. 5), and is delayed in all low contrast trials, but additionally delayed for color decisions. The differential delay persists for several hundred milliseconds and is also visible in the peak latency of the P3 ERP component on go trials (stimulus evaluation) and the mean response times for go trials (behavior).

grounded memory access in the visual system and other sensory, motor, and affective systems, as well as their roles in language comprehension.

Acknowledgments

This research was supported by the Center for Research in Language Postdoctoral NIDCD Fellowship T32DC000041-12 and the Institute for Neural Computation Postdoctoral NIMH Fellowship T32MH020002-13 to B.D.A., and NICHD Grant 22614 to M.K.

Conflict of interest

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

References

Allport, D.A., 1985. Distributed memory, modular subsystems and dysphasia. In: Newman, S.K., Epstein, R. (Eds.), Current Perspectives in Dysphasia. Churchill Livingstone, NY, pp. 32–60.

Amsel, B.D., 2011. Tracking real-time neural activation of conceptual knowledge using single-trial event-related potentials. Neuropsychologia 49 (5), 970–983.

Amsel, B.D., Urbach, T.P., Kutas, M., 2013. Alive and grasping: stable and rapid semantic access to an object category but not object graspability. NeuroImage 77, 1–13.

Avidan, G., Harel, M., Hendler, T., Ben-Bashat, D., Zohary, E., Malach, R., 2002. Contrast sensitivity in human visual areas and its relationship to object recognition. J. Neurophysiol. 87 (6).

Barros-Loscertales, A., Gonzalez, J., Pulvermüller, F., Ventura-Campos, N., Carlos Bustamante, J., Costumero, V., et al., 2012. Reading salt activates gustatory brain regions: fMRI evidence for semantic grounding in a novel sensory modality. Cereb. Cortex 22 (11).

Barsalou, L.W., 1999. Perceptual symbol systems. Behav. Brain Sci. 22 (4), 577–660.

Barsalou, L.W., 2008. Grounded cognition. Annu. Rev. Psychol. 59, 617–645.

Bedny, M., Caramazza, A., 2011. Perception, action, and word meanings in the human brain: the case from action verbs. Ann. N. Y. Acad. Sci. 1224, 81–95.

Bedny, M., Caramazza, A., Pascual-Leone, A., Saxe, R., 2012. Typical neural representations of action verbs develop without vision. Cereb. Cortex 22 (2), 286–293.

Benjamini, Y., Krieger, A.M., Yekutieli, D., 2006. Adaptive linear step-up procedures that control the false discovery rate. Biometrika 93 (3), 491–507.

Boulenger, V., Roy, A.C., Paulignan, Y., Deprez, V., Jeannerod, M., Nazir, T.A., 2006. Cross-talk between language processes and overt motor behavior in the first 200 msec of processing. J. Cogn. Neurosci. 18 (10), 1607–1615.

Bouvier, S.E., Engel, S.A., 2006. Behavioral deficits and cortical damage loci in cerebral achromatopsia. Cereb. Cortex 16 (2), 183–191.

Brang, D., Hubbard, E.M., Coulson, S., Huang, M.X., Song, T., Ramachandran, V.S., 2010. Magnetoencepalography reveals early activation of V4 in grapheme color synesthesia. NeuroImage 53 (1), 268–274.

Caramazza, A., Hillis, A.E., Rapp, B.C., Romani, C., 1990. The multiple semantics hypothesis — multiple confusions. Cogn. Neuropsychol. 7 (3), 161–189.

Chan, A.M., Baker, J.M., Eskandar, E., Schomer, D., Ulbert, I., Marinkovic, K., et al., 2011. First-pass selectivity for semantic categories in human anteroventral temporal lobe. J. Neurosci. 31 (49), 18119–18129.

Chao, L.L., Martin, A., 1999. Cortical regions associated with perceiving, naming, and knowing about colors. J. Cogn. Neurosci. 11 (1), 25–35.

Chatterjee, A., 2010. Disembodying cognition. Lang. Cogn. 2 (1), 79–116.

Dils, A.T., Boroditsky, L., 2010. Visual motion aftereffect from understanding motion language. Proc. Natl. Acad. Sci. U. S. A. 107 (37), 16396–16400.

Ganis, G., Kutas, M., Sereno, M.I., 1996. The search for "common sense": an electrophysiological study of the comprehension of words and pictures in reading. J. Cogn. Neurosci. 8 (2), 89–106.

Gonzalez, J., Barros-Loscertales, A., Pulvermüller, F., Meseguer, V., Sanjuan, A., Belloch, V., et al., 2006. Reading cinnamon activates olfactory brain regions. NeuroImage 32 (2), 906–912.

Grill-Spector, K., Malach, R., 2004. The human visual cortex. Annu. Rev. Neurosci. 27.

Groppe, D.M., Urbach, T.P., Kutas, M., 2011. Mass univariate analysis of event-related brain potentials/fields II: simulation studies. Psychophysiology 48 (12), 1726–1737.

Gullick, M.M., Mitra, P., Coch, D., 2013. Imagining the truth and the moon: an electrophysiological study of abstract and concrete word processing. Psychophysiology 50 (5), 431–440.

Halgren, E., Wang, C.M., Schomer, D.L., Knake, S., Marinkovic, K., Wu, J.L., et al., 2006. Processing stages underlying word recognition in the anteroventral temporal lobe. NeuroImage 30 (4).

Hauk, O., Tschentscher, N., 2013. The body of evidence: what can neuroscience tell us about embodied semantics? Front. Psychol. 4.

Hauk, O., Coutout, C., Holden, A., Chen, Y., 2012. The time-course of single-word reading: evidence from fast behavioral and brain responses. NeuroImage 60 (2), 1462–1477.

Hoenig, K., Sim, E., Bochev, V., Herrnberger, B., Kiefer, M., 2008. Conceptual flexibility in the human brain: dynamic recruitment of semantic maps from visual, motor, and motion-related areas. J. Cogn. Neurosci. 20 (10), 1799–1814.

Hsu, N.S., Kraemer, D.J.M., Oliver, R.T., Schlichting, M.L., Thompson-Schill, S.L., 2011. Color, context, and cognitive style: variations in color knowledge retrieval as a function of task and subject variables. J. Cogn. Neurosci. 23 (9), 2544–2557.

Hubbard, E.M., Manohar, S., Ramachandran, V.S., 2006. Contrast affects the strength of synesthetic colors. Cortex 42 (2).

Kaschak, M.P., Madden, C.J., Therriault, D.J., Yaxley, R.H., Aveyard, M., Blanchard, A.A., et al., 2005. Perception of motion affects language processing. Cognition 94 (3), B79–B89.

Kastner, S., Pinsk, M.A., De Weerd, P., Desimone, R., Ungerleider, L.G., 1999. Increased activity in human visual cortex during directed attention in the absence of visual stimulation. Neuron 22 (4).

Landauer, T.K., Dumais, S.T., 1997. A solution to Plato's problem: the latent semantic analysis theory of acquisition, induction, and representation of knowledge. Psychol. Rev. 104 (2), 211–240.

Lavric, A., Pizzagalli, D.A., Forstmeier, S., 2004. When 'go' and 'nogo' are equally frequent: ERP components and cortical tomography. Eur. J. Neurosci. 20 (9), 2483–2488.

Legge, G.E., Ahn, S.J., Klitz, T.S., Luebker, A., 1997. Psychophysics of reading—XVI. the visual span in normal and low vision. Vis. Res. 37 (14), 1999–2010.

Levelt, W.J.M., 1989. Speaking: From Intention to Articulation. , MIT Press, Cambridge, MA.

Mahon, B.Z., Caramazza, A., 2008. A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. J. Physiol. Paris 102 (1–3), 59–70.

Mahon, B.Z., Anzellotti, S., Schwarzbach, J., Zampini, M., Caramazza, A., 2009. Category-specific organization in the human brain does not require visual experience. Neuron 63 (3), 397–405.

Martin, A., 2007. The representation of object concepts in the brain. Annu. Rev. Psychol. 58, 25–45.

McRae, K., Cree, G.S., Seidenberg, M.S., McNorgan, C., 2005. Semantic feature production norms for a large set of living and nonliving things. Behav. Res. Methods 37 (4), 547–559.

Mechelli, A., Humphreys, G.W., Mayall, K., Olson, A., Price, C.J., 2000. Differential effects of word length and visual contrast in the fusiform and lingual gyri during reading. Proc. R. Soc. B Biol. Sci. 267 (1455).

Meteyard, L., Bahrami, B., Vigliocco, G., 2007. Motion detection and motion verbs — language affects low-level visual perception. Psychol. Sci. 18 (11), 1007–1013.

Meteyard, L., Zokaei, N., Bahrami, B., Vigliocco, G., 2008. Visual motion interferes with lexical decision on motion words. Curr. Biol. 18 (17), R732–R733.

Miceli, G., Fouch, E., Capasso, R., Shelton, J.R., Tomaiuolo, F., Caramazza, A., 2001. The dissociation of color from form and function knowledge. Nat. Neurosci. 4 (6), 662–667.

Moscoso del Prado, M., Hauk, O., Pulvermüller, F., 2006. Category specificity in the processing of color-related and form-related words: an ERP study. NeuroImage 29 (1), 29–37.

Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9 (1).

Oram, M.W., Xiao, D., Dritschel, B., Payne, K.R., 2002. The temporal resolution of neural codes: does response latency have a unique role? Philos. Trans. R. Soc. B Biol. Sci. 357 (1424), 987–1001.

Paivio, A., 1986. Mental Representation: A Dual-coding Approach. , Oxford University Press, New York.

Pecher, D., Zeelenberg, R., Barsalou, L.W., 2003. Verifying different-modality properties for concepts produces switching costs. Psychol. Sci. 14 (2).

Pulvermüller, F., 2013. Semantic embodiment, disembodiment or misembodiment? In search of meaning in modules and neuron circuits. Brain Lang. 127 (1), 86–103.

Pulvermuller, F., Hauk, O., Nikulin, V.V., Ilmoniemi, R.J., 2005. Functional links between motor and language systems. Eur. J. Neurosci. 21 (3).

Pulvermüller, F., Shtyrov, Y., Hauk, O., 2009. Understanding in an instant: neurophysiological evidence for mechanistic language circuits in the brain. Brain Lang. 110 (2), 81–94.

Pylyshyn, Z.W., 1984. Computation and Cognition. , MIT Press, Cambridge, MA.

Ramachandran, V.S., Hubbard, E.M., 2001. Synaesthesia: a window into perception, thought and language. J. Conscious. Stud. 8 (12), 3e34.

Riddoch, M.J., Humphreys, G.W., Coltheart, M., Funnell, E., 1988. Semantic systems of system — neuropsychological evidence re-examined. Cogn. Neuropsychol. 5 (1), 3–25.

Rodriguez-Fornells, A., Schmitt, B.M., Kutas, M., Munte, T.F., 2002. Electrophysiological estimates of the time course of semantic and phonological encoding during listening and naming. Neuropsychologia 40 (7), 778–787.

Rouw, R., Scholte, H.S., 2007. Increased structural connectivity in grapheme-color synesthesia. Nature Neuroscience 10, 792–797.

Rugg, M.D., Thompson-Schill, S.L., 2013. Moving forward with fMRI data. Perspect. Psychol. Sci. 8 (1), 84–87.

Sasaki, K., Gemba, H., Nambu, A., Matsuzaki, R., 1993. No-go activity in the frontal association cortex of human-subjects. Neurosci. Res. 18 (3), 249–252.

Schmitt, B.M., Munte, T.F., Kutas, M., 2000. Electrophysiological estimates of the time course of semantic and phonological encoding during implicit picture naming. Psychophysiology 37 (4), 473–484.

Schmitt, B.M., Schiltz, K., Zaake, W., Kutas, M., Munte, T.F., 2001. An electrophysiological analysis of the time course of conceptual and syntactic encoding during tacit picture naming. J. Cogn. Neurosci. 13 (4), 510–522.

Simmons, W.K., Ramjee, V., Beauchamp, M.S., McRae, K., Martin, A., Barsalou, L.W., 2007. A common neural substrate for perceiving and knowing about color. Neuropsychologia 45 (12), 2802–2810.

Thorpe, S., Fize, D., Marlot, C., 1996. Speed of processing in the human visual system. Nature 381 (6582), 520–522.

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 (7), 1915–1926.

Tootell, R.B.H., Reppas, J.B., Kwong, K.K., Malach, R., Born, R.T., Brady, T.J., et al., 1995. Functional-analysis of human mt and related visual cortical areas using magnetic-resonance-imaging. J. Neurosci. 15 (4).

van Dantzig, S., Pecher, D., Zeelenberg, R., Barsalou, L.W., 2008. Perceptual processing affects conceptual processing. Cogn. Sci. 32 (3), 579–590.

VanRullen, R., Thorpe, S.J., 2001. The time course of visual processing: from early perception to decision-making. J. Cogn. Neurosci. 13 (4), 454–461.

Warrington, E.K., Shallice, T., 1984. Category specific semantic impairments. Brain 107, 829–854 (SEP).

Weiss, P.H., Fink, G.R., 2009. Grapheme-colour synaesthetes show increased grey matter volumes of parietal and fusiform cortex. Brain 132, 65–70.

West, W.C., Holcomb, P.J., 2000. Imaginal, semantic, and surface-level processing of concrete and abstract words: an electrophysiological investigation. J. Cogn. Neurosci. 12 (6), 1024–1037.