Title
Colour and Stability in Embodied Representations

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 27(27)

ISSN
1069-7977

Author
Connell, Louise

Publication Date
2005

Peer reviewed

Colour and Stability in Embodied Representations

Louise Connell (louise.connell@northumbria.ac.uk)
Division of Psychology, Northumbria University
Newcastle upon Tyne, NE1 8ST, UK

Abstract

Traditional views of cognition have assumed that thought involves the representation and manipulation of discrete, amodal symbols, while embodied theories of cognition hold that thought is grounded in the same neural systems that govern sensation, perception and action. This paper examines whether implicit perceptual information on object colour is represented during sentence comprehension even though doing so does not necessarily facilitate task performance. After reading a sentence that implied a particular colour for a given object, participants were presented with a picture of that object that either matched or mismatched the implied colour. When asked if the pictured object was mentioned in the preceding sentence, results showed that people's responses were faster (but less accurate) when the colours *mismatched* than when they *matched*. A distinction between stable and unstable embodied representations is proposed to allow embodied theories to account for these findings, and discussed with reference to future directions in cognitive modelling.

Introduction

Imagine a person sitting at a desk wondering whether to have a sandwich or soup for lunch. This is an everyday cognitive feat, and yet is fraught with many unresolved issues. How do we represent a sandwich that is not actually in front of us at the time? How do we represent the notion of soup without a specific flavour or colour? The ability to form and utilise conceptual knowledge is central to human cognitive life, and how we manage to do this is a key question in cognitive psychology and cognitive science.

Amodal Representations

Many traditional theories of conceptual thought have assumed that concepts are discrete, amodal representations (e.g., Katz & Fodor, 1963; Kintsch & van Dijk, 1978; Fodor, 1975; Pylyshyn, 1984). In such an amodal representation, our hypothetical bowl of soup could be represented simply as SOUP, and we would not have to concern ourselves with properties such as flavour or colour unless we have that information available to represent (e.g. FLAVOUR:TOMATO or COLOUR:RED). Such amodal theories of representation have been popular in cognitive science for their computational convenience, and are a cornerstone of symbolic cognitive modelling in domains from language comprehension (Kintsch, 2001) and plausibility judgement (Connell & Keane, in sub.), to analogical reasoning (Falkenhainer, Forbus & Gentner, 1989; Keane, Ledgeway & Duff, 1994) and more general cognitive architectures (Anderson, 1993; Newell, 1990). There are significant advantages in using an amodal approach in cognitive models, as the representational shortcut offered by amodal symbols facilitates the modeling and testing of large-scale, complex, cognitive tasks.

However, criticism of amodal representations has become increasingly frequent of late (Barsalou, 1999; Glenberg & Kachak, 2002; Johnson-Laird, 1983; Pecher & Zwaan, 2005). The main criticisms concern the interaction between the external and internal world: how perception can map to arbitrary conceptual symbols (the transduction problem: Barsalou, 1999), and how conceptual symbols can map back to the real world (the grounding problem: Harnad, 1990). Even if the bowl of savoury liquid in front of us can be related to the arbitrary symbol SOUP, and if the symbol SOUP can be related to other similar bowls of liquid, the means by which this perceptual↔conceptual translation takes place remains unknown. Without any reference to the outside world, amodal symbols can have no meaning (Searle, 1980), and this has led some researchers to propose more perceptually-grounded embodied theories of conceptual thought.

Embodied Representations

A growing body of empirical work has recently emerged in support of embodied representations of concepts. Neuroimaging studies have shown how sensorimotor areas in the brain are activated during language processing (Carpenter et al., 1999; Pulvermüller, 1999). In addition, it has been shown that "low-level" sensorimotor representations play a role in "high-level" cognitive processes such as language comprehension and memory retrieval (Glenberg & Kaschak, 2002; Richardson et al., 2003; Solomon & Barsalou, 2001; Stanfield & Zwaan, 2001; Zwaan, Stanfield & Yaxley, 2002). Importantly, these studies employed implicit tasks such as recognition and naming, demonstrating that perceptual information is activated even though doing so does not facilitate task performance. For example, Stanfield and Zwaan (2001) presented people with sentences that mention objects with implied orientation (e.g., "Rick put the pencil in the cup" or "Rick put the pencil in the drawer"), followed by a picture of an object (e.g. a pencil). People were faster to verify that a pencil had been mentioned in the sentence when it was pictured in the orientation implied by the sentence (i.e. pictured vertically for the cup sentence, pictured horizontally for the drawer sentence). This finding is incompatible with amodal theories of conceptual knowledge, which cannot account for why the orientation of the pencil is represented when it is not specified in the sentence. However, Stanfield and Zwaan's findings can be explained if participants construct an embodied representation of the sentence (e.g. a sensorimotor simulation of placing a pencil in a cup/drawer), as this would include implied information about the pencil's orientation.

Although such evidence is often presented as underscoring the endemic problems of amodal theories of representation, there is currently no single theory of embodied cognition that can integrate their findings. Rather, theories of embodied representation share certain characteristics and assumptions (see Wilson, 2002) and consider perception and action as central to higher cognition. One of the most influential of these theories is Barsalou's (1999) Perceptual Symbol Systems. According to this theory, concepts are essentially partial recordings of the neural activation that arises during perceptual and motor experiences. These recordings (or "perceptual symbols") can later be re-enacted as a perceptual simulation of that concept.

For example, to represent the concept of *soup*, neural systems for vision, action, touch, taste, smell etc. partially reproduce our previous experiences of *soup*. The perceptual symbols activated for *soup* may include visual information of liquid in a bowl, sensorimotor information of eating hot savoury liquid with a spoon, etc. Barsalou (1999) thus argues that Perceptual Symbol Systems theory avoids transduction and grounding problems by assuming that conceptual representations are based on the same systems that are used for perception and action.

Representing Colour Information

Colour representation is a key aspect of perceptual information that has not received the same attention in the embodiment debate as other visual object attributes such as shape, size and orientation. However, some studies have already indicated that colour may not be represented as a context-free amodal attribute (such as COLOUR: X). For example, when asked to compare the colour *grey* to *black* and *white*, Medin and Shoben (1988) found that people considered *grey* to be more similar to *white* in the context of *hair*, but more similar to *black* in the context of *clouds*. Similarly, Halff, Ortony, and Anderson (1976) found that people represented the colour *red* differently for *hair*, *wine*, *flag*, *brick*, and *blood*, considering the colour of a *red* flag to be more similar to a *red* light than a *red* wine. Rather than using a single amodal symbol such as COLOUR:RED across concepts, people appeared to use different perceptual, embodied representations.

However, a colour such as *red* is not a constant parameter: differing wavelengths of light give rise to shades of *red* that vary in hue, saturation, and luminosity. In that respect, it is perhaps not surprising that people consider *red* wine differently to *red* bricks. These findings are not necessarily incompatible with amodal representations if one argues that people are retrieving knowledge of context-specific, subordinate shades of the colour category *red* (e.g. COLOR:WINE-RED, COLOR:BRICK-RED) rather than a generic superordinate COLOR:RED. Presenting people with explicit colour terms is not a suitable paradigm for distinguishing whether amodal or embodied symbols are activated in representing colour information.

The Current Study

In an amodal representation, the sentence "John looked at the steak" could be represented by a proposition such as LOOK(JOHN, STEAK). But what about the colour of the steak: is it red (raw) or brown (cooked)? This information is not contained in the sentence and so we do not have a colour attribute available for STEAK. An embodied representation of this sentence would also lack this colour information, as simulating an individual *steak* does not require simulating its colour. [1]

On the other hand, what if the sentence read "John looked at the steak on his plate."? As before, an amodal representation does not encode any colour information as it only represents the explicit information in the sentence, such as LOOK(JOHN,STEAK[LOCATION:ON_PLATE]). However, although the sentence contains no explicit colour information, the placing of a steak on a plate *implies* that the steak is cooked and therefore has a brown colour. An embodied representation of a sentence would contain such implied perceptual information by simulating an individual *plate* and *steak* and then specialising the colour of the steak to *brown*.

The study reported in this paper examines whether implicit perceptual information on object colour is represented during sentence comprehension. Such a finding would be incompatible with traditional amodal theories of representation and cognition, and would lend support to the embodied view of mental representation. In this study, participants are asked to perform an implicit recognition task that tests whether perceptual information is activated even though doing so does not facilitate task performance. Results are then related back to amodal and embodied theories of representation, and are discussed with a view to the future of both theories within cognitive science.

Experiment

This experiment presents participants with short sentences followed by a picture, and asks them to indicate whether the pictured object was mentioned in the sentence. For test items, the pictured object was always mentioned in the preceding sentence but the object was shown in one of two picture conditions: *matching* the colour implied by the sentence or *mismatching* the colour implied in the sentence . For example, the sentence "John looked at the steak on his plate." was followed by a picture of a brown steak in the match condition, but by a red steak in the mismatch condition (note that each colour used in the match and mismatch picture conditions was a valid colour representation of that particular object: raw steaks are *red*, cooked steaks are *brown*). In addition, sentences had two different forms, each of which implied one of the object colours. Figure 1 shows a sample of the sentences versions and pictures used in each condition. Thus, the basic design crossed two sentence versions (version 1, version 2) with two picture conditions (match, mismatch).

[1] A default colour could be represented for *steak* in either an amodal or embodied representation, but this is not important for our present purposes as it does not distinguish the two theories.

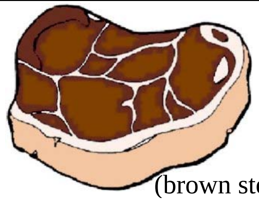
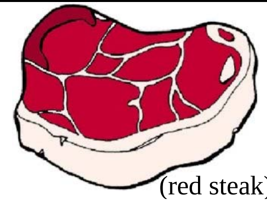
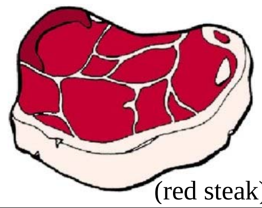
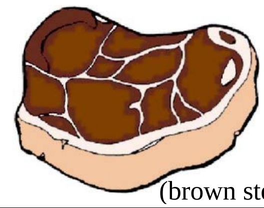
Sentence	Picture Condition	
	Match	Mismatch
John looked at the steak on his plate.	 (brown steak)	 (red steak)
John looked at the steak in the butcher’s window.	 (red steak)	 (brown steak)

Figure 1: Sample sentences used in experiment, showing pictures used in match and mismatch conditions

There are two important aspects of this design that differ from earlier studies such as Halff et al. (1976) that were not concerned with the amodal/embodied distinction. First, as previously discussed, colour information is not explicitly stated but rather is implied by the sentential context. Second, rather than keeping colour constant and varying the possible objects, the object has been kept constant and its possible colours varied. This allows amodal and embodied representations to be teased apart because of their differing predictions regarding participant response latencies

In amodal representations, both sentence versions shown in Figure 1 would be represented without any colour information. Thus, the traditional amodal view would predict that people would respond equally quickly and accurately in the match and mismatch conditions because they are simply confirming that the pictured object (e.g. STEAK) was mentioned in the preceding sentence (e.g. LOOK(JOHN,STEAK[LOCATION:ON_PLATE]). The embodied view, however, makes very different predictions. In embodied representations, both sentence versions would be represented with the implied colour encoded as part of the simulation for *steak* – i.e. simulating a steak on a plate would involve specializing the steak colour to the appropriate *brown*, while simulating a steak in a butcher’s window would involve specializing the steak colour to the appropriate *red*. Thus, the embodied view would predict that people will be faster and more accurate (confirming that the pictured object was mentioned in the sentence) in the match condition than in the mismatch condition.

Method

Materials. Forty-four pictures were created for use in this experiment. Of these, twenty-four were test items (forming pairs of pictures) and twenty were fillers (unrelated standalone pictures). Many of the pictures came from popular clipart packages but some were created by the author. All pictures were coloured naturalistically by sampling shades from photographs of the relevant objects, and contained only one predominant colour (e.g. Figure 1’s red steak predominantly contains shades of *red*). Each pair of test pictures was identical except for the colours used. All pictures were resized to a maximum of 250 pixel height (approx. 6.9cm onscreen) and 350 pixel width (approx. 9.7cm onscreen).

It was important that picture recognition would not be affected by the canonicality or view specificity of the pictures (see Stanfield and Zwaan, 2001), and so each picture was pretested to meet these requirements. Twenty-two participants were presented onscreen with an object name followed by a picture, and had to indicate whether the picture matched the name. Each pair of test pictures (e.g. *red* / *brown* steak) was separated to form two groups of items and participants were randomly assigned to one of the groups. Filler pictures were seen by both groups and were presented after semantically unrelated words, thus requiring a “no” response. All test pictures were presented after their object name and required a “yes” response. After indicating their yes/no response, participants were asked to rate the general quality of the picture on a scale from 1 (poor quality) to 7 (good quality). All pictures used as test items in this experiment met the following criteria: the median response time for each item was < 1250ms, there was no significant difference in response times between the pictures in each test pair (all *p*s > 0.2), and each item received a median quality rating of at least 4 out of 7. The sole criterion for filler items was that each received a median quality rating of at least 4 out of 7.

Forty-four sentences were constructed to accompany the pictures in this experiment. Of these, twenty-four were test items (naming an object featured in a test picture) and twenty were fillers (naming objects not featured in either the test or filler pictures). The test sentences thus formed pairs, with each member of a pair implying a different colour for the same object. Filler sentences all contained at least one concrete noun. In order to ensure that the test sentences actually implied the intended colour for the object, another pretest (using 24 new participants) was conducted. Each pair of test sentences was separated to form two groups of items and participants were randomly assigned to one of the groups. Each sentence was presented along with two pictures of an object mentioned in the sentence (i.e. both matching and mismatching pictures) and participants were asked to choose, from four forced-choice alternatives, whether a) the first picture best matched the sentence, b) the second picture best matched the sentence, c) both pictures equally match the sentence, or d) neither picture matches the sentence. All test items used in this experiment met the criterion of having the picture from the matching condition chosen at least 50% of the time.

Design. Test items were divided into four groups so that each group featured one of four sentence-picture combinations: version1-match, version1-mismatch, version2-match, version2-mismatch. Each group contained equal numbers of match and mismatch test items, and the various colours featured in test pictures were distributed approximately evenly across groups. Participants were assigned randomly to one of the groups. Thus, the experiment was a 2 (sentence version: version1, version2) × 2 (picture condition: match, mismatch) × 4 (group) design, with sentence version and picture condition as within-participants variables and group as a between-participants variable.

Participants. Sixty native speakers of English from Northumbria University (not used in pretests) were paid a nominal sum for participation in this experiment.

Procedure. Testing took place on portable computers running Presentation software. Participants read instructions describing the experiment and instructing them to read each sentence and then to decide if the pictured object had been mentioned in the preceding sentence. Participants were asked to respond as quickly as possible as their response time was being measured, and to read every sentence carefully as their comprehension would be tested at various points during the experiment. Each trial began with a left-aligned vertically-centred fixation cross presented for 1000ms, followed by presentation of a sentence. When participants pressed the space bar to indicate comprehension another fixation cross was displayed centrally onscreen for 500ms, followed by a picture. Participants had to decide if the pictured object had appeared in the preceding sentence and indicate their decision by pressing the key labeled "yes" (the comma key) or the key labeled "no" (the full stop key). In half of all filler trials, a comprehension question (relating to the filler sentence) appeared after the picture decision. Participants were required to answer an equal number of "yes" and "no" comprehension questions. A blank screen was displayed for 500ms as an inter-stimulus break between trials. The entire procedure took approximately 10 minutes.

Table 1: Mean response times with standard deviations (in ms) and accuracy rates (%) for match and mismatch picture conditions.

Picture Condition	Mean RT (SD)	Accuracy
Match	1328 (577)	93.6%
Mismatch	1190 (542)	70.2%

Results & Discussion

Two participants that answered <50% of the comprehension questions correctly were eliminated from the analysis. One further participant was also excluded for failing to respond with valid keystrokes. All responses <300ms and >3000ms were considered outliers and dropped from the analysis, as were any responses more than two standard deviations away from a participant's mean in the relevant condition. Altogether, 9% of the data was excluded in this way.

Results were not wholly consistent with either the amodal or embodied views of representation, but were partially consistent with the embodied view. Table 1 shows the mean correct response times and accuracy for the match and mismatch picture conditions. Analyses of variance were run on the data by participants and by items, and interactions involving the group variable are not reported due to their lack of theoretical importance. Against both amodal and embodied predictions, people responded more quickly when the picture colour *mismatched* (*M*=1190ms, *SD*=542ms) the object colour implied by the sentence than when it *matched* (*M*=1328ms, *SD*=577ms), with analysis significant by participants, $F_1(1, 41)=7.845$, $MSE=0.082$, $p<0.01$; $F_2(1, 31)=2.156$, $MSE=0.070$, $p=0.15$. The interaction of sentence version × picture condition was significant by participants, $F_1(1, 41)=6.212$, $MSE=0.103$, $p<0.05$; $F_2<1$. Accuracy (responding correctly that the pictured object was mentioned in the preceding sentence) was significantly higher when the picture colour matched the implied colour (93.6%) than when it mismatched (70.2%), in line with embodied predictions, $F_1(1, 53)=56.056$, $p<0.0001$; $F_2(1, 32)=14.018$, $p<0.001$]. The interaction of sentence version × picture condition was not significant (*F*s<1).

The profile of response times was much slower than other experiments using a similar picture decision paradigm (e.g. Stanfield & Zwaan, 2001; Zwaan et al., 2003), inviting the possibility that these slower response latencies reflect very different processing and task strategies. However, even when the fastest quartile of response times per condition is taken[2], the same pattern of response times emerges with the mismatch picture condition (687ms) being faster than the match condition (734ms). It should be noted that the other studies mentioned used black and white line drawings while the present study uses coloured line drawings, and this extra complexity may influence processing times.

So what do these results mean for the contrasting positions of amodal and embodied representation? The predictions of the amodal view (that implied sentence colour would have no effect on people's response speed and accuracy) were not borne out by the data. The main prediction of the embodied view (that matching sentence and picture colours would facilitate faster responses) was not supported either. Indeed, the exact inverse was found with mismatching colours being faster than matching. Only the accuracy prediction made by the embodied view (that matching colours would make people more accurate in confirming that the pictured object was mentioned in the sentence) found support in the results. The match condition

[2] Above a minimum threshold of 300ms, response times at the 25th percentile were 977ms (match) and 890ms (mismatch).

was significantly more accurate than the mismatch condition, as also found by Zwaan et al. (2002). However, it is interesting to note that the pattern of errors in this study does not conform to speed-acuracy tradeoff: accuracy improves as responses become faster, reaching 97.6% (match) and 91.8% (mismatch) in the fastest quartile [2]. Together, these findings suggest that many of the criticisms levelled at amodal theories are accurate: simple propositional representations of sentences cannot account for the results found in this study. However, it is also apparent that current theories of embodied representation, such as Barsalou's (1999) Perceptual Symbol Systems, are at present ill-equipped to explain these findings.

While it seems counterintuitive that people's response times were facilitated by mismatching rather than matching colours, Naor-Raz, Tarr and Kersten (2003) reported a similar pattern of results in a modified Stroop task. They showed participants names of concrete objects displayed in different colours (e.g. "banana" in yellow or purple text) and measured naming times for the text colour. They found that people were faster when text colour *mismatched* the named object's colour (e.g. "banana" in purple text) than when text and object colour *matched* (e.g. "banana" in yellow text). Naor-Raz et al.'s results suggest that participants found it easier to ignore incongruent information about object colour than to ignore congruent colour information. The same rationale can be applied to the results reported here. When the colour of the pictured object is different to that implied by the sentence, people can compare the pictured object to their mental representation and confirm rapidly that yes, the object was indeed mentioned in the preceding sentence (ignoring differences in other properties). Conversely, when the colour of the pictured object is the same as that implied by the sentence, people compare the pictured object to their mental representation and confirm that yes, the object was indeed mentioned in the preceding sentence and yes, it even matches in colour (ignoring differences in other properties). In other words, responses in the mismatching condition may be faster than those in the matching condition because the implied colour information in the sentence representation is easier to ignore. There is some benefit to the extra processing time consumed in the matching condition: attending to the implied colour information makes participants more likely to confirm accurately that the pictured object was mentioned in the preceding sentence.

But why should implied colour information produce such different behaviour to implied orientation (Stanfield & Zwaan, 2001) or shape (Zwaan et al., 2002) information? While Kaschak et al. (2005) also found mismatch facilitation when examining object motion, their explanation – a visual stimulus "ties up" neural motion mechanisms and hinders ability to process a simultaneously-presented auditory sentence that describes motion in the same direction – cannot explain results in this study's sequential presentation paradigm. One other possible explanation is that colour, unlike other object attributes such as shape, size, and orientation, does not have as stable a specialisation in an embodied representation. Studies of visual memory have suggested that colour is not as salient as other properties that determine the configuration of a scene (e.g., object presence, position, or shape) and hence that colour is encoded with *less stability* in scene representations (Aginsky & Tarr, 2000; Vandenbeld & Rensink, 2003). If an embodied sentence representation involves a perceptual simulation of visual information, then it is possible that the colour of an object (as a non-configurational property) is not represented as a stable specialisation. The instability of colour information in an embodied representation could be a contributing factor to the overall slower response times and increased error rates observed in this experiment compared to experiments concerned with more stable, configurational shape and orientation information. More importantly, the instability of embodied colour information could also account for faster response times in the mismatch versus match condition. This notion will be explored further in the general discussion.

General Discussion

In this work, amodal and embodied theories of representation are contrasted with a study examining the representation of implied colour information. Results showed that perceptual colour information is activated during sentence comprehension even though doing so does not facilitate task performance. People responded more accurately when the colour of a pictured object matched the colour implied by the previous sentence. This finding is in line with embodied theories and is incompatible with amodal, propositional theories which hold that implied perceptual information is not represented. In addition, people were found to respond more slowly when the colour of a pictured object matched the colour implied by the previous sentence. This finding is also incompatible with amodal theories and is contrary to that predicted by current embodied theories which hold that matching implied information should facilitate faster responses. So, this paper proposes a distinction between stable and unstable embodied representations as explanation for the results.

According to Perceptual Symbol Systems theory (Barsalou, 1999), the neural systems that represent colour in perception also represent object colours (as perceptual simulations). For example, thinking about a *red car* involves activating the partial recordings of neural activation that arose during previous perceptual experiences with cars (i.e. a perceptual simulation of *car*) and then specialising the colour to *red*. In short, perceptual properties of objects – visual, aural, tactile, etc. – are represented by specialising the perceptual simulation of the relevant object. In this paper, the focus of the experiment was the representation of visual information. When information that determines the configuration of a visual scene (shape, size, orientation, etc.) is implied by a sentence, it is represented in a perceptual simulation as a *stable* specialisation of the relevant object. This stability makes implied information difficult to ignore when a mismatch occurs between a pictured object (e.g. a vertical pencil) and the preceding sentence (e.g. "Rick put the pencil in the drawer") (Stanfield & Zwaan, 2001). So, for stable embodied representations, people are slower to respond in a mismatching condition than a matching condition. Conversely, when information that is not salient to the configuration of a visual scene (such as colour) is implied by a sentence, it is represented in a perceptual simulation as an *unstable* specialisation of the relevant object. This instability makes the implied information relatively easy to ignore when a mismatch occurs between a pictured object (e.g. a brown steak) and the preceding sentence (e.g. "John looked at the steak in the butcher's window"). Thus, for unstable embodied representations, people are faster to respond in a mismatching condition than a matching condition, as reported in the present study. Further research on this topic (Connell, in prep.) examines whether this stable / unstable distinction is specific to embodied representations of familiar objects.

Nevertheless, amodal theories of representation are widely used in cognitive science and cognitive modelling, often to valuable effect. For this reason, it is important to observe that embodied theories of representation are not necessarily incompatible with the idea that cognitive tasks can be described in terms of computational processing. Many cognitive models are not functionally dependent on discrete symbols, and their views of cognitive processing would be valid irrespective of the form of the mental representations actually being processed. As evidence grows for embodied mental representations, we gain a better understanding of exactly what information people represent and how they use it across a variety of cognitive tasks. Future cognitive models that address such tasks should examine ways to integrate the embodied evidence and obviate explicit commitment to ungrounded symbols. In this way, cognitive modelling and embodied research could mutually inform theories of cognitive processing.

Acknowledgments

This work was funded by the Division of Psychology in Northumbria University. Many thanks to Dermot Lynott for helpful comments and discussion, and to Ben Singleton and Darren Dunning for data collection.

References

Aginsky, V., & Tarr, M. J. (2000). How are different properties of a scene encoded in visual memory? *Visual Cognition, 7*, 147-162.

Anderson, J. (1993). *Rules of the Mind.* Hillsdale, NJ: Erlbaum.

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

Carpenter, P. A., Just, M. A., Keller, T. A., Eddy, W. F., Thulborn, K. R. (1999). Time course of fMRI-activation in language and spatial networks during sentence comprehension. *NeuroImage, 10,* 216-224.

Connell, L. (in prep.). Embodied representations of colour. *Manuscript in preparation.*

Connell, L., & Keane, M. T. (in sub.). A model of plausibility. *Manuscript in submission.*

Falkenhainer, B., Forbus, K., & Gentner, D. (1989). The Structure-Mapping engine: algorithm and examples. *Artificial Intelligence, 41*, 1-63.

Fodor, J. (1975). *The language of thought.* Cambridge: Harvard University Press.

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

Halff, H. M., Ortony, A. & Anderson, R. C. (1976) A context-sensitive representation of word meanings. *Memory & Cognition 4,* 378–83.

Harnad, S. (1990). The symbol grounding problem. *Physica D, 42*, 335-346.

Johnson-Laird, P. N. (1983). *Mental models.* Cambridge, MA: Harvard University Press.

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

Katz, J. J. & Fodor, J. A, (1963). The structure of a Semantic Theory, *Language, 39*, 170-210.

Keane, M. T., Ledgeway, T., & Duff, S. (1994). Constraints on analogical mapping: A comparison of three models. *Cognitive Science, 18*, 387-438.

Kintsch, W. (2001). Predication. *Cognitive Science, 25*, 173-202.

Kintsch, W. & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*, 363-394.

Medin, D., & Shoben, E. (1988). Context and structure in conceptual combinations. *Cognitive Psychology, 20,* 158-190.

Naor-Raz, G., Tarr, M. J., & Kersten, D. (2003). Is color an intrinsic property of object representation? *Perception, 32*, 667-680.

Newell, A. (1990). *Unified theories of cognition.* Cambridge, MA: Harvard University Press.

Pecher, D., & Zwaan, R. A. (2005). Introduction to grounding cognition. In D. Pecher & R. A. Zwaan (Eds.), *Grounding cognition: the role of perception and action in memory, language, and thinking.* Cambridge: Cambridge University Press.

Pulvermüller, F. (1999). Words in the brain's language. *Behavioral & Brain Sciences, 22*, 253–336.

Pylyshyn, Z. W. (1984). *Computation and cognition.* Cambridge, MA: MIT Press.

Richardson, D. C., Spivey, M. J., Barsalou, L. W., & McRae, K. (2003). Spatial representations activated during real-time comprehension of verbs. *Cognitive Science, 27*, 767-780.

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences, 3,* 417–24.

Solomon, K. O., & Barsalou, L. W. (2001). Representing properties locally. *Cognitive Psychology 43*, 129–169.

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

Vandenbeld, L. A., & Rensink, R. A. The decay characteristics of size, color, and shape information in visual short-term memory. *Journal of Vision, 3*. 682.

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

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