BRAIN RESEARCH 1183 (2007) 109–123

available at www.sciencedirect.com

www.elsevier.com/locate/brainres

BRAIN RESEARCH

Research Report

Immediate integration of novel meanings: N400 support for an embodied view of language comprehension

Dorothee J. Chwilla, Herman H.J. Kolk, Constance T.W.M. Vissers*

Nijmegen Institute for Cognition and Information, Radboud University, Nijmegen, The Netherlands

ARTICLE INFO

Article history:
Accepted 2 September 2007
Available online 19 September 2007

Keywords:
Language
ERPs
N400
Theories of meaning
Integration model
Abstract symbol theories of meaning
Latent semantic analysis
Embodied theories of meaning

ABSTRACT

A substantial part of language understanding depends on our previous experiences, but part of it consists of the creation of new meanings. Such new meanings cannot be retrieved from memory but still have to be constructed. The goals of this article were: first, to explore the nature of new meaning creation, and second, to test abstract symbol theories against embodied theories of meaning. We presented context-setting sentences followed by a test sentence to which ERPs were recorded that described a novel sensible or novel senseless situation (e.g., "The boys searched for *branches*/bushes [*sensible*/senseless] with which they went drumming…"). Novel sensible contexts that were not associatively nor semantically related were matched to novel senseless contexts in terms of familiarity and semantic similarity by Latent Semantic Analysis (LSA). Abstract symbol theories like LSA cannot explain facilitation for novel sensible situations, whereas the embodied theory of Glenberg and Robertson [Glenberg, A.M., Robertson, D.A., 2000. Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. Journal of Memory and Language, 43, 379–401.] in which meaning is grounded in perception and action can account for facilitation. Experiment 1 revealed an N400 effect in a sensibility judgment task. Experiment 2 demonstrated that this effect generalizes to a situation in which participants read for comprehension. Our findings support the following conclusions: First, participants can establish new meanings not stored in memory. Second, this is the first ERP study that shows that N400 is sensitive to new meanings and that these are created immediately – that is, in the same time frame as associative and semantic relations. Third, our N400 effects support embodied theories of meaning and challenge abstract symbol theories that can only discover meaningfulness by consulting stored symbolic knowledge.

1. Introduction

Meaning in cognitive psychology has received very little attention compared to syntax. This is surprising because meaning more than structural factors guides human behavior. The search for meaning or common sense is omnipresent (e.g., Hess et al., 1995). Imagine you see a man smashing a window of a car with a hammer. In everyday life people spontaneously come up with several ideas to make sense out of other's behavior. Likely, one of the first ideas that based on our

* *Corresponding author*. Nijmegen Institute for Cognition and Information, Radboud University, P.O. Box 9104, 6500 HE Nijmegen, The Netherlands. Fax: +31 24 3616066.
E-mail addresses: chwilla@nici.ru.nl (D.J. Chwilla), kolk@nici.ru.nl (H.H.J. Kolk), c.vissers@nici.ru.nl (C.T.W.M. Vissers).

0006-8993/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.brainres.2007.09.014

knowledge of the world[1] comes to mind is that the man is a burglar who tries to steal valuables from the car. Suppose, however, that you get to know that the man is the owner of the car. Then people readily make up different stories, for instance, that the man lost his keys and desperately needs something inside of the car. Other possible stories would be that the man is furious and tries to prevent that his girlfriend leaves him by car or that the man is furious that his car does not work because due to this he will miss his airplane.

This example reveals two important aspects about human behavior: First, that people on the basis of past experiences have certain expectancy patterns that are helpful to interpret behavior. In cognitive psychology, these expectancy patterns are referred to as *scripts* or schemata. A good example is the restaurant script (Schank and Abelson, 1977) which involves action sequences like sitting at a table, ordering food from a menu, and drinking wine. For instance, on entering a restaurant we expect the waiter to give us a menu but not the check. Daily life for the most part consists of such scripts that help organize the world such that it becomes manageable.

A second impressive aspect (revealed by the example) is that people are creative in accounting for other's behavior – that is, they do not stick to explanations in terms of familiar (script) situations but invent new explanations if needed. What is meaningful depends on the circumstances, for example drying one's feet with socks is weird when somebody has a towel but not if one has not. What is striking is the apparent ease with which people seem to generate new "hypotheses" about behavior. This ability is remarkable because new meanings are not stored in semantic memory and thus still have to be constructed. This in contrast with old meanings such as scripts that represent familiar combinations in long-term memory, and therefore are readily available based on our knowledge of the world.

The goal of this article is to shed light on the nature of new meaning creation by using event-related brain potentials (ERPs). Before we will describe our experimental approach, we will briefly sketch theories of meaning. The classical theories of meaning in the field of linguistics and psycholinguistics are abstract symbol theories (e.g., Anderson et al., 1997; Collins and Loftus, 1975; Fodor, 1983; Masson, 1995; Ratcliff and McKoon, 1988). A main assumption of these theories is that meaning arises from the syntactic combination of abstract, amodal (i.e., nonperceptual) symbols that are arbitrarily related to entities in the real world. For example, according to one of the most frequently cited theories of Collins and Loftus (1975), meaning arises from the pattern of relations among nodes in a network. Every node in the network corresponds to an undefined word, and the set of nodes to which a particular node is connected corresponds to the words in the dictionary definition.

This conception gave rise to different types of representation, such as feature lists, schemata, semantic nets, and connectionism. According to abstract symbol theories words have fixed meanings (e.g., Katz and Fodor, 1963). Specifically, the meaning of a word can be captured by defining a set of individually necessary and jointly sufficient features. On this view, for each word there exists a mental list of essential features. For example, to identify whether an object is an elephant one has to check whether the object possesses all essential features of an elephant (trunk, big animal, grey color, large ears, herbivore, etc.). Recently mathematical high-dimensional models of meaning (i.e., Latent Semantic Analysis [LSA] of Landauer and Dumais (1997) and Hyperspace Analogue to Language (HAL) of Burgess and Lund (2000)) have been presented as a new variant of the classical abstract symbol theories. The two high-dimensional models claim that the meaning of a word is its vector representation in a high-dimensional space, and these vectors are very similar to the abstract symbols used in the classical theories of meaning. A major strength of the classical approach to meaning is that abstract amodal symbols have been very successful in representing several kinds of knowledge (i.e., types and tokens, propositions and abstract concepts) and to combine symbols productively.

On the negative side abstract symbol theories have been criticized on the following grounds: First, it has been shown that the meanings of words are not fixed but graded and dependent on the context (see e.g., Barsalou, 1993; Lakoff, 1987). Second, it has been called into question how meaning can arise from abstract symbols that are not grounded in perception or action (referred to as the grounding problem). As forcefully argued by Searle in his famous Chinese Room example (1980) manipulation of abstract amodal symbols produces more abstract symbols not meaning. To illustrate this point in a different way: Picture yourself in a restaurant in a country like Japan where you do not speak the native language. You are entering a restaurant and a waiter brings you a Japanese menu. The only other source of information you have at your disposal is a Japanese lexicon. No matter how many abstract symbols you will look up in the lexicon it will not give you a clue about what the symbols on the menu stand for.

An alternative approach that more recently gained popularity in linguistics and cognitive psychology in general is the embodied framework (e.g., Barsalou, 1999; Jackendoff, 1987; Lakoff, 1987; Pecher and Zwaan, 2005). What the different theories have in common is that they are perceptual theories of cognition. A main assumption of the embodied approach is that mental processes such as thinking or language understanding are based on the physical interactions that people have with their environment. As pointed out by Pecher and Zwaan (2005), "Rather than viewing the body as a support system for a mind that needs to be fueled and transported, they view the mind as a support system that facilitates the functioning of the body. By shifting the basis for mental behavior toward the body, these researchers assume that mental processes are supported by the same processes that are used for physical interactions, that is, for perception and action. Cognitive structures develop from perception and action". On this view, meaning does not arise from the combination of abstract symbols into a representational structure. Instead meaning is based on our current (physical) interactions or previous experiences of interactions with objects in different kinds of environments. Current or past body by environment interactions guide people in how to think about, that is, *simulate* the perceptual and action details required by a situation.

Recently, the embodied framework has been applied to the field of language comprehension (see for example: Glenberg, 1997; Glenberg et al., 2005; Glenberg and Robertson, 1999; 2000; Yeh and Barsalou, 2006; Zwaan and Madden, 2005; Zwaan and

[1] World knowledge refers to our notions of what kinds of objects and events occur in the world around us.

Radvansky, 1998). As outlined above the starting point is that the structure of the body is very important in that it determines the range of effective actions. Here the term *affordances* plays a crucial role. Following Gibson (1979), embodied researchers define affordances as the actions that are suggested by a particular object (e.g., Borghi, 2005). These affordances probably arise based on learning. So they are the result of previous experiences and interactions with objects. But the range of possible actions suggested by an object also includes new actions that have not been previously performed. Borghi (2005) has argued that in the case of simple actions – as opposed to complex actions – visual input and object knowledge assist subjects to extract affordances automatically. On this view concepts can be perceived as patterns *of potential action* that support people in extracting affordances (Glenberg, 1997). We will come back to this below.

In this article, the focus is on the embodied view of language comprehension as described in Glenberg and Robertson (1999, 2000). They developed the Indexical Hypothesis to apply the embodied theory to the field of language comprehension. According to this hypothesis, understanding a sentence such as, "*Jareb stood on the chair to change a light bulb*," involves three processing steps. The first step is to index phrases to actual objects or analogical perceptual symbols (Barsalou, 1999) representing the objects. Thus, the noun phrase "the chair" may be taken to refer to an actual chair in the perceiver's environment or indexed to a prototypical representation of a chair that retains perceptual information; that is, a perceptual symbol. The second step is to use the indexed object or perceptual symbol to derive affordances.[2] For example, a chair by an adult is typically used to sit on, but can also be used to stand on to reach an object that otherwise would be out of reach. The third step is to *mesh* the affordances guided by the syntax of the sentence. This means that the language user combines the actions suggested by the chair and the actor so that Jareb is on the chair rather than, say, under the chair. Glenberg and Robertson emphasize that the "meshing" depends on an individual's experience and learning history. For instance, for a child sitting under a chair (e.g., hiding) could be a quite familiar experience while for an adult this normally does not work. A sentence is meaningful to the language user to the extent that (s)he can mesh the objects and activities as directed by the sentence. One main advantage of the embodied approach is that because the model is based on perception it brings us closer to a solution of the grounding problem (Searle, 1980; see also Harnad, 1990). On the negative side, embodied theories have been criticized as being too vague and it is clear that a further specification (in case of the Indexical Hypothesis of the different processing steps) is needed before the merits of this approach to language comprehension can be fully assessed.

The general goal of this article was to explore the nature of the process of new meaning creation. The more specific aim was to test abstract symbol theories against embodied theories of meaning. The approach used in this article was similar to that of Glenberg and Robertson (2000). Therefore, we will describe the latter research in more detail. Glenberg and Robertson (2000) tested abstract symbol theories, in particular Latent Semantic Analysis (LSA), against embodied theories of meaning, in particular their Indexical Hypothesis, by presenting sentences to subjects that either described a sensible novel context or a nonsense novel context. For example, a context-setting sentence like "Marissa forgot to bring her pillow on her camping trip" could be followed by a novel sensible sentence (e.g., "As a substitute for her *pillow*, she filled up an old sweater with *leaves*.") or by a novel but senseless sentence (e.g., "As a substitute for her *pillow*, she filled up an old sweater with *water*."). Via careful pretests of the materials, they verified that the novel contexts were indeed new. In particular, Glenberg and Robertson demonstrated by using Latent Semantic Analysis that in more than 35 000 texts, a central concept of each sentence (e.g., "pillow") appears in contexts that are orthogonal to the contexts in which the distinguishing concepts (e.g., "water" and "leaves") appear. The two classes of theories make different predictions with respect to the processing of novel meanings. Abstract symbol theories of meaning cannot explain that subjects can make sense out of novel sensible situations. The reason why these theories cannot explain an improvement in performance for novel sensible vs. novel insensible situations is that the only way they can discover meaningfulness is by consulting stored symbolic knowledge. However, this possibility was precluded by controlling LSA values. In contrast, embodied theories of meaning can explain that humans are able to make sense out of new information not stored in long-term memory. On this view, people conceptualize specific situations, including those in which language is involved, by simulating themselves as full-bodied beings in these events. According to the Indexical Hypothesis, meaning for both familiar situations and for novel situations evolves from meshing the affordances of the objects and activities as described by a sentence or discourse. What is meaningful and what is not depends on our knowledge about the possibilities and limitations of the human body and individual experiences. Glenberg and Robertson (2000) demonstrated that participants judged new sensible sentences as more meaningful than new senseless sentences. Furthermore, imageability ratings correlated highly with the sensibility ratings, which fits with the view that subjects derive meaningfulness by how well affordances can be meshed. Their data thus supported embodied theories of meaning and challenged abstract symbol theories of meaning.

In the present article, we used ERPs to shed light on the nature of the mental simulation process that gives rise to facilitation for novel sensible meanings compared to novel senseless meanings. This simulation process plays a central role in embodied theories of language comprehension. One could argue that this hypothetical simulation process does not represent an alternative to a symbolic way of processing but only an additional option some kind of strategy that the language system could use only under certain circumstances. One major question of the present studies was to determine, whether the mental simulation process to novel meanings reflects a more automatic process – that is built-in in the language system or reflects a more strategic and controlled

[2] We would like to point out that words – in contrast with objects – do not have affordances in that they support interactions of the body with the environment. On the other hand any object can have a very large number of affordances. Which of the affordances are simulated is partly determined by syntax.

process. In the former case, facilitation effects for novel meanings inform us about the architecture of the language system whereas in the latter case these effects would not inform us about the organization of language but instead would reflect an additional function. The nature of the mental simulation process was determined by testing for sensibility effects when the sensibility of the critical sentences was in the focus of attention (Experiment 1) and when the sensibility was not in the focus of attention in that subjects did nothing else than what they do in everyday life that is read for comprehension (Experiment 2).

We investigated this question by comparing the ERP response to novel sensible meanings with that to novel senseless meanings. ERPs provide a continuous record of brain activity and therefore allow us to track the time course of meaning integration on a millisecond-by-millisecond basis. The focus is on one language-relevant ERP component, *the N400* that has been shown to be highly sensitive to semantic processing. The most frequently cited view is that N400 is an index of meaning integration (e.g., Chwilla et al., 2000; Friederici, 1995; Holcomb, 1993; Van Berkum et al., 1999). Modulations in N400 amplitude reflect the ease with which different kinds of knowledge (e.g., semantic, conceptual, perceptual) are integrated into context, be this a single word, a sentence or discourse. Some results, in particular the finding of ERP masked priming effects (e.g., Kiefer, 2002; Stenberg et al., 2000) and of N400 effects in the attentional blink paradigm (e.g., Rolke et al., 2001), have been presented as evidence against the integration view. Specifically, these findings have been taken to indicate that the N400 does not *solely* reflect semantic integration processes. However, while these results show that N400 in addition to more controlled processes also reflects automatic processes,[3] this does not imply that these N400 effects – as proposed by these authors – were caused by spreading activation (see for reviews on semantic priming, Hutchison, 2003; McNamara, 2005; and for a discussion of automatic priming models see Chwilla and Kolk, 2002). As pointed out by McNamara (2005) with respect to the masked N400 effects the possibility exists that semantic integration processes operate even when primes – or context in general – are presented outside conscious awareness. This hypothesis is consistent with recent demonstrations of list composition effects on unconscious semantic priming (e.g., Bodner and Masson, 2003). Relevant for our present purposes is that Kintsch (1988) proposed a model of text comprehension that includes automatic semantic integration processes of this kind. Empirical support for automatic integration comes from the demonstration of N400 backward priming effects at very short intervals (Chwilla et al., 1998). Thus although in the literature we often see a coupling between automatic processes and spreading activation, the finding of automatic priming per se can be accounted for by integrative mechanisms and, thus, is compatible with the N400 integration view.

The N400 has been shown to be an effective tool to track semantic integration processes online. Recent studies have shown that different kinds of world knowledge, in particular well-known facts (e.g., like “Dutch trains are *yellow*…” [see: Hagoort et al., 2004]) and scripts (three words forming a familiar scenario like “Director – Bribe – Dismissal” [see Chwilla and Kolk, 2005]) are integrated as fast as semantic knowledge. These EPR studies show that *familiar* types of world knowledge (old meanings) are integrated immediately into the higher-order meaning representation of the context. Important for the present argument is that all kinds of knowledge that up to now have been shown to elicit N400 effects, represent familiar knowledge stored in long-term memory. We propose that world knowledge based on previous experiences is represented in a large semantic network in long-term memory. The structure of this network can be made visible by means of LSA. In case of isolated words, it is also possible to determine the association strength or category membership between items. However, there may also be knowledge that is not stored in the semantic network but that nevertheless affects the ease with which the meaning of a sentence is captured. This information is stored in the knowledge that we have about our body. It is this knowledge that we refer to as “embodied”. To access and retrieve embodied knowledge, we have to project ourselves into a situation and simulate the perceptual and action details required by that situation. In this article, we tested whether the N400 is also sensitive to integration difficulties that arise from embodied knowledge. To ensure that the integration difficulties originate from embodied knowledge, it has to be ruled out that the integration problem may arise from other familiar kinds of knowledge. In the present studies, we controlled for this by matching novel sensible contexts and novel senseless contexts in terms of familiarity and semantic similarity by LSA. If, despite of this matching, N400 effects to novel sensible vs. novel senseless contexts are observed then these effects must reflect embodied knowledge. With respect to the direction of the N400 effect we predicted that N400 amplitude would be reduced for novel affordable actions (e.g., to drum with branches) compared to novel non-affordable actions (e.g., to drum with bushes).

The N400 also allows us to determine the nature of the mental simulation process underlying the construction of novel meanings not stored in long-term memory. In particular, we propose that an immediate N400 effect to novel sensible meanings compared to novel insensible meanings – that is an effect within the typical time frame of the classical N400 effect to associative and/or semantic relations – would be compatible with a more automatic simulation process. In contrast, a delay in timing of the N400 effect would be compatible with the view that the simulation process is of a more controlled nature.

As far as we know nobody yet has probed for N400 effects to novel meanings. Of course, a necessary condition for a determination of the processing nature of the mental simulation process by the N400 is that a reliable N400 effect is obtained to novel sensible meanings. A first experiment, therefore, was conducted to test whether novel sensible contexts compared to novel senseless contexts indeed elicit an N400 effect. To facilitate a comparison of the ERP results with the behavioral results in the literature, the procedure and task used in Experiment 1 were very similar to that used by Glenberg and Robertson (2000). That is, in Experiment 1, participants after reading the critical sentence to which their

[3] The term automatic processes is used in the traditional sense referring to processes that have a quick onset, proceed without intention or awareness, and produce benefits but not costs.

brain activity was recorded indicated on a scale from 1 to 7 how meaningful the sentence given the context had been (7 indicates very meaningful and 1 very senseless).

2. Experiment 1

2.1. Results

2.1.1. Delayed sensibility judgment task

An analysis of the sensibility ratings indicated that novel sensible situations were judged as more meaningful than novel insensible situations [5.6 vs. 2.4; $F(1,29)=999{,}91$, $p<.001$].

2.1.2. Event-related potentials

The electrode montage for Experiments 1 and 2 is presented in Fig. 1. The grand mean waveforms time locked to the critical words for all midline sites and a representative set of lateral sites are displayed in Fig. 2. As Fig. 2 shows, the most distinguishing feature of the waveforms is a negative-going wave that starts around 300 ms and peaks at about 400 ms. The negative wave resembled the standard N400 in terms of the morphology, the timing and the scalp distribution. Therefore, in the following, we refer to this negative wave as an N400. The N400 was widely distributed across the scalp and was followed by a positive-going shift peaking at about 900 ms largest at central/parietal sites. Inspection of Fig. 2 suggests that at central/posterior midline sites (see Cz, and Pz) and bilateral posterior sites (see P3, P3p, P4 and RTP), N400 amplitude was smaller for novel sensible than novel insensible critical words. No other ERP differences between conditions were observed.

2.1.3. Statistical analyses

N400 window (300–500 ms). For the midline sites, a main effect of sensibility was found, $F(1,29)=4.66$, $p<.05$. This effect confirmed that mean N400 amplitude was smaller for novel sensible critical words (0.85 μV) than for novel insensible critical words (−0.05 μV). The sensibility by site interaction did not approach significance, $F(4,116)=2.47$, $p=.110$, indicating that the N400 sensibility effect was broadly distributed across the midline.

For the lateral sites, no main effect of sensibility was found, $F<1.6$. However, a two-way sensibility by site interaction, $F(4,116)=3.06$, $p<.05$, and a three-way sensibility by ROI by site interaction were obtained, $F(4,116)=4.93$, $p<.01$. Follow-up analyses disclosed that these interactions reflected that the N400 sensibility effects were restricted to posterior ROIs. Specifically, follow-up analyses for the single sites revealed N400 effects for two posterior sites of the left hemisphere (P3 and P3p: p-values$<.03$). The sensibility by hemisphere interaction did not approach significance, $F<1$. Planned comparisons for the two hemispheres did not disclose main effects of sensibility, both F values<2. For the left hemisphere, there was a trend towards a sensibility by site interaction, $F(10,290)=3.01$, $p=.07$, reflecting N400 effects at P3 and P3p and a trend towards an N400 effect at LTP ($p<.06$) but not at other sites. Also for the right hemisphere the sensibility by site interaction, $F(10,290)=3.30$, $p=.058$, approached significance. The latter interaction reflected a trend towards an N400 effect at two posterior sites of

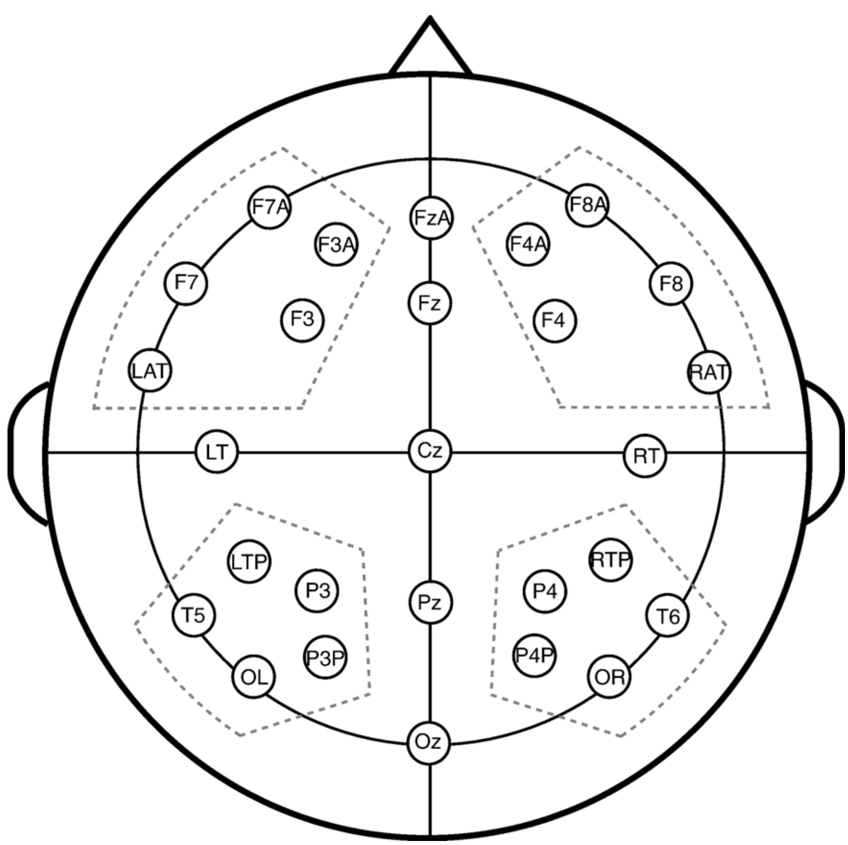

Fig. 1 – The electrode configuration of Experiments 1 and 2.

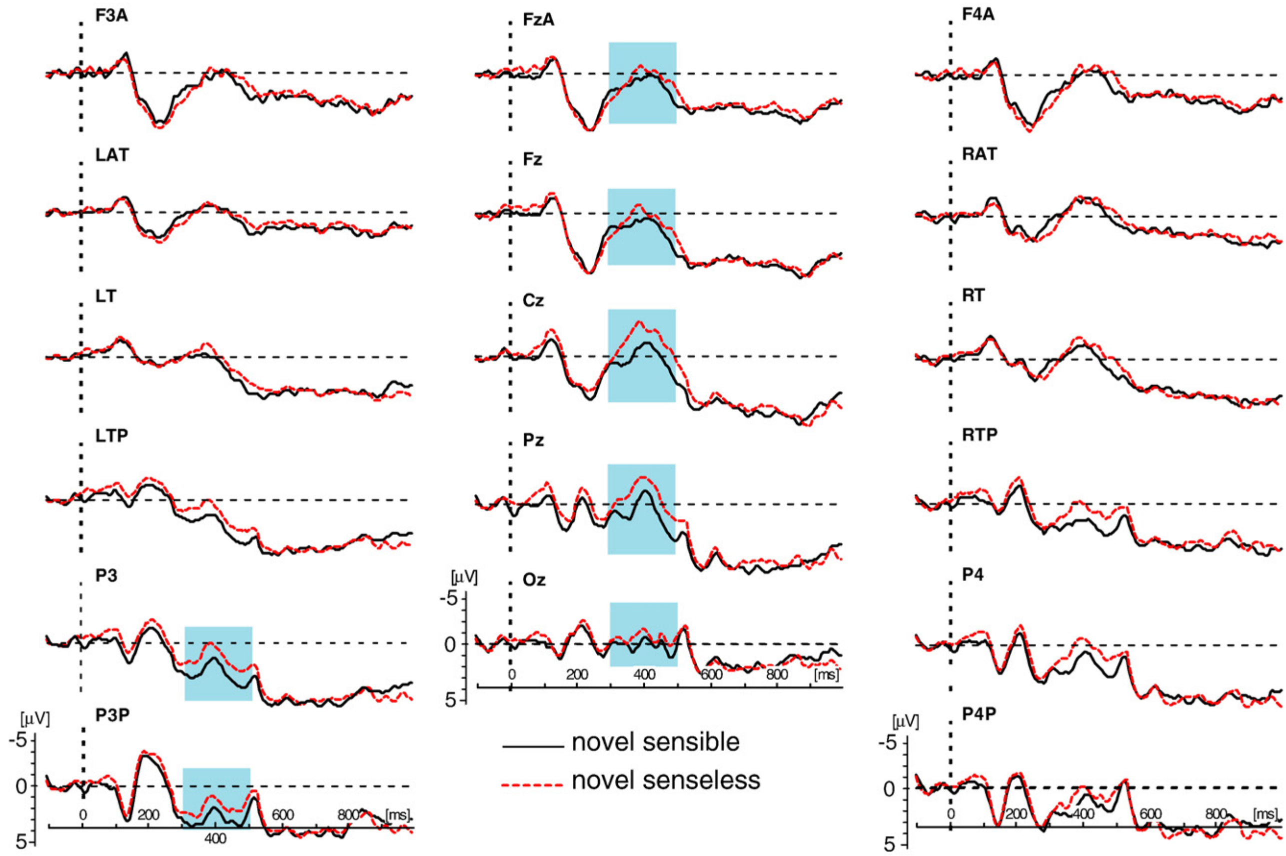

Fig. 2 – Experiment 1. The grand mean ERP waveforms for the sensibility judgment task for a representative set of sites, time locked to the onset of the critical word (time point zero), relative to a 100-ms pre-critical word baseline. Negativity is plotted upward. Shaded areas indicate the time periods and the set of electrodes that showed significant N400 effects.

the right hemisphere (RTP and P4: $p<.07$). A supplementary N400 analysis for the 400 to 500 ms window (see Experimental procedures) for the lateral sites revealed significant N400 effects for the following sites of the left hemisphere (P3, P3p and OL: p values $<.05$) and more importantly also for two posterior sites of the right hemisphere (RTP and P4: p values $<.05$).

The analyses on N400 peak amplitude yielded a trend towards an effect of sensibility for the midline sites, $F(1,29)=4.06$, $p=.053$, in the absence of a sensibility by site interaction, $F(4,116)=2.16$, $p=.13$. The trend reflected that N400 peak amplitude tended to be smaller to novel sensible critical words (−2.93 μV) than to novel senseless critical words (−3.96 μV). To explore whether significant effects on N400 peak amplitude were present at any site, t-tests were carried out for the single sites. These tests disclosed a sensibility effect on N400 peak amplitude at Cz and Pz (p values $<.03$). For the lateral sites no sensibility effect, $F(1,29)=2.50$, $p=.125$, or other interactions with sensibility were found, neither in the overall analysis nor in the planned comparisons for the left and the right hemispheres (all F values <1.5).

2.2. *Discussion*

Consistent with previous results of Glenberg and Robertson (2000), participants judged novel sensible contexts as more meaningful than novel senseless contexts. Note that in a speeded reaction time (RT) version of Experiment 1, no facilitation in RT or error rates for novel sensible contexts was observed. The error data of the pilot study, however, showed that participants are quite accurate in distinguishing novel sensible from novel insensible contexts as reflected by that mean error rates in both conditions were lower than 12% (for details see Experimental procedures: Reaction time pilot study).

Let us now turn to the ERP data. As described in the introduction, N400 effects have been reported to several kinds of familiar knowledge including associative, semantic and world knowledge (well-known facts and script knowledge). As far as we know, nobody has yet tested for N400 effects for novel sensible meanings. Therefore, the main question investigated in Experiment 1 was whether novel sensible contexts not stored in long-term memory produce an N400 effect or not. The main result of Experiment 1 was that novel sensible contexts that were matched in terms of familiarity/semantic similarity by LSA to novel senseless contexts elicited an N400 effect. In other words, we succeeded in constructing novel sensible stories, which according to LSA were not more familiar in terms of association strength or semantic relatedness than the control stories, but nevertheless yielded an N400 effect. The ERP data are of theoretical relevance in that the N400 sensibility effect cannot be explained by abstract symbol theories of meaning (e.g., Anderson et al., 1997; Collins and Loftus, 1975; Fodor, 1983; Katz and Fodor, 1963; Masson, 1995; Ratcliff and McKoon, 1988). Because the critical words in the sentences were neither associatively nor semantically related according to these models, no facilitation should occur for novel sensible contexts. The results show that participants can establish new meanings that are not stored in memory, and that the N400 is sensitive to such new meanings. The finding of an N400 effect to novel sensible contexts can be accounted for by the embodied theory of language comprehension of Glenberg and Robertson (1999, 2000). Therefore, the ERP data provide further support for an embodied theory of reading and call into question abstract symbol theories of meaning.

The ERP data allow us to track the time course of new meaning creation in real time. The N400 effect to novel sensible contexts (in the following termed *the N400 sensibility effect*) occurred in the same time frame as the typical N400 effect to simple associative or semantic relations. It showed an onset at about 300 ms and peaked around 400 ms (e.g., Van Herten et al., 2006). The similarity in time course is taken to indicate that new meanings are constructed and integrated as fast as familiar (old) meanings. The immediacy of the N400 sensibility effect to novel meanings is striking because participants could not base their decisions on previous experiences with the objects but had to rely on on-the-fly imaginative simulations. This finding is in line with the view that simulations of actions are obligatory and the hypothesis that embodiment functions automatically and interactively in people's use and understanding of linguistic meaning (see Gibbs, 2005).

Were there other indications for differences between the N400 sensibility effect and the classical N400 effect? No, both the waveshape and the scalp distribution of the N400 effect to novel sensible meanings were similar to that of the classical N400 effect. In particular, the N400 sensibility effect resembled the N400 effect to associative and/or semantic relations in that it was maximal at central/posterior midline sites and bilateral posterior sites. Based on the literature on creative language like metaphor or joke comprehension (e.g., Coulson and Wu, 2005), one might have predicted a right-hemisphere preponderance for the N400 sensibility effect for new meanings. However, no such asymmetry was observed. One possibility is that the bilateral distribution of the N400 effect in Experiment 1 is due to the use of an additional judgment task as opposed to reading tasks employed in most other studies.

To sum up, the main result of Experiment 1 was that an N400 sensibility effect to novel sensible contexts was obtained. Importantly, this N400 sensibility effect was similar to the standard N400 effect in terms of its timing and scalp distribution. These results support the view that novel sensible words are immediately accessed and integrated into the ongoing context.

3. Experiment 2

3.1. *Introduction*

After having established in Experiment 1 that novel sensible meanings give rise to an N400 effect, Experiment 2 was conducted to determine the nature of the N400 effect. A possible critique on Experiment 1, the studies of Glenberg and Robertson (2000) and some other studies that have been presented as evidence for the embodied framework (see e.g., Glenberg and Kaschak, 2002; Stanfield and Zwaan, 2001) is, that they exploited a judgment task. In case of the present Experiment 1 and the studies of Glenberg and Robertson (2000), the task directed the participant's attention at the sensibility of the test sentence. It is clear that this task does not comprise part of the reading process. Therefore, it could be argued that the behavioral and the online N400 sensibility effects as well arise from task-induced strategies and do not reflect processes of reading. On this view, immediate integration of novel meanings may only occur when processing of this type of information is in the focus of attention or overtly requested. In the latter case, these facilitation effects for novel meanings would not inform us about the architecture of the language comprehension system but instead would reflect an additional function.

What would be needed is a measure of language comprehension that does not involve metacognitive reflection. If we could demonstrate that novel meanings are integrated immediately when subjects do nothing else than read for comprehension then we could rule out the hypothesis that the N400 effects to novel sensible contexts were caused by task-specific strategies. The N400 is exactly such a measure. It has been shown that N400 effects occur in reading tasks across several contexts (isolated words, word triplets, sentences, discourse). The goal of Experiment 2, therefore, was to investigate if the N400 effect observed in the sensibility judgment task in Experiment 1 generalizes to a reading situation. Experiment 2 was a replication of Experiment 1 except that no additional task was given. Participants in Experiment 2 did nothing else than read for comprehension. The predictions were as follows: If the N400 effect to novel sensible contexts in Experiment 1 reflects an additional function (i.e., strategies induced by the judgment task) then the N400 effect should disappear in a reading task. In contrast, if the N400 sensibility effect reflects reading processes, then the effect should also show up in Experiment 2. Most important for our present purposes is that the occurrence of an immediate N400 sensibility effect in Experiment 2, would support first, embodied theories of language comprehension, and second, the view that the mental simulation process represents an alternative to a symbolic way of processing and not an additional function that the language system only uses in particular circumstances.

3.2. *Results*

3.2.1. *Event-related potentials*

The mean waveforms for novel sensible and novel senseless critical words for all midline sites and a representative set of bilateral sites are presented in Fig. 3. The overall form of the ERPs was very similar to that of Experiment 1. Again, a negative wave that resembled the N400 in terms of its waveform and timing was obtained at the midline and bilateral sites. Inspection of Fig. 3 suggests that novel sensible critical words again elicited an N400 effect compared to novel senseless critical words. The only difference with Experiment 1 seemed to be the scalp distribution of the effect: the N400-

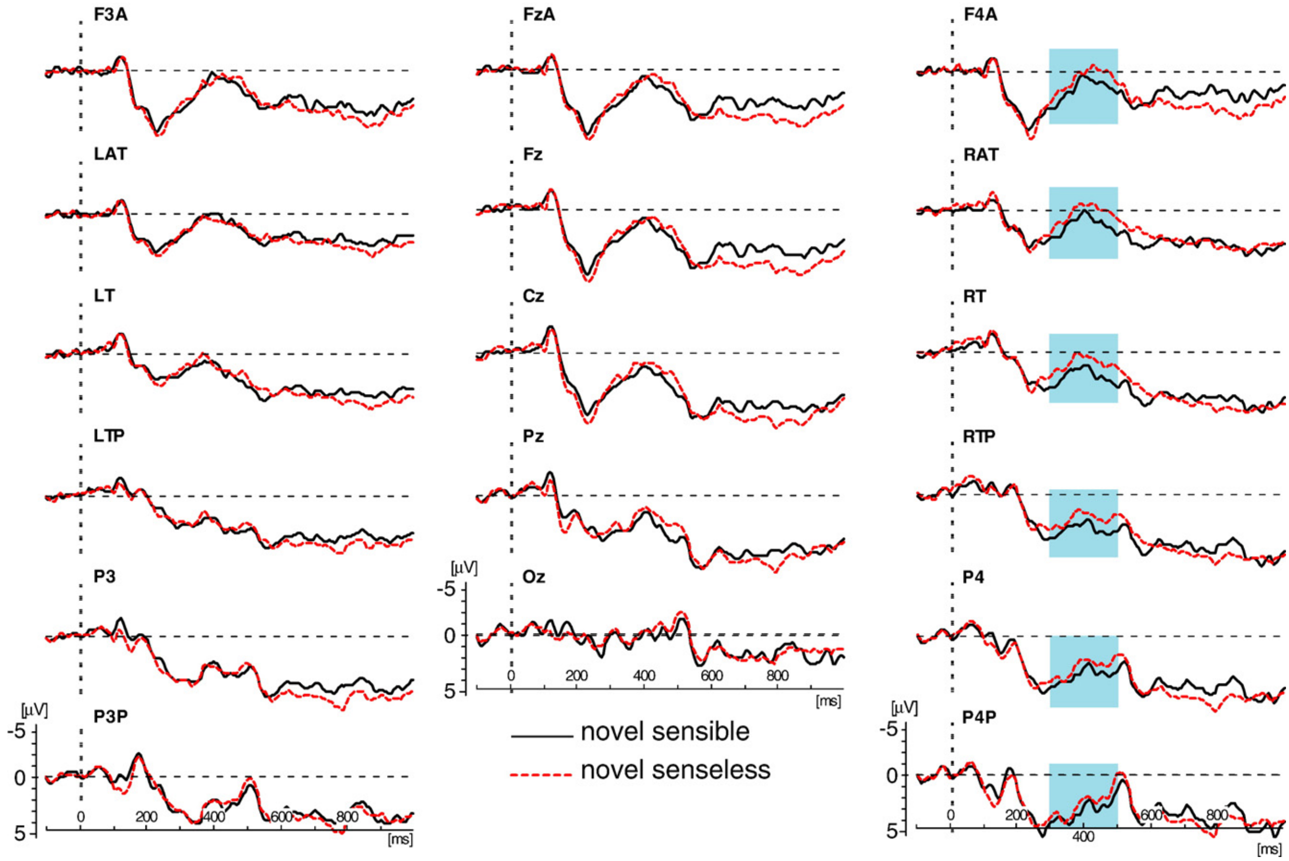

Fig. 3 – Experiment 2. The grand mean ERP waveforms during reading for a representative set of sites, time locked to the onset of the critical word (time point zero), relative to a 100-ms pre-critical word baseline. Negativity is plotted upward. Shaded areas indicate the time periods and the set of electrodes that showed significant N400 effects.

like effect in the reading task was not largest at central–parietal midline sites but most pronounced over the right hemisphere including anterior and posterior regions. The N400 was followed by a positivity that at some sites (frontocentral midline sites and bilateral posterior sites) seemed to be larger, that is, more positive for novel sensible than novel senseless contexts.[4]

3.2.2. *Statistical analyses*

N400 window (300–500 ms). The analyses on mean N400 amplitude for the midline sites did not yield an effect of sensibility, $F<1.6$, or an interaction of this factor with site, $F<1$. For the lateral sites there was a trend towards an effect of sensibility, $F(1,34)=2.94$, $p<.10$. More importantly, a significant two-way sensibility by hemisphere interaction was obtained, $F(1,34)=10.72$, $p<.01$. Separate analyses for the two hemispheres revealed a main effect of sensibility for the right hemisphere, $F(1,34)=5.44$, $p<.03$. No sensibility by site interaction was present, $F<1$, indicating that the N400 sensibility effect was widely distributed including anterior and posterior areas of the right hemisphere. In contrast, for the left hemisphere, there was no indication for an N400 sensibility effect, in that no effect of sensibility, $F<1$, or sensibility by site interaction was found, $F<1$.

The analyses on N400 peak amplitude for the midline sites did not yield a main effect of sensibility, $F<2.2$, or an interaction of this factor with site, $F<2.7$, $p=.113$. For the lateral sites, a main effect of sensibility was present, $F(1,34)=4.30$, $p<.05$. The main effect disclosed that mean N400 peak amplitude was significantly smaller for novel sensible critical words (−1.33 µV) than for novel senseless critical words (−2.03 µV). The sensibility by hemisphere interaction was not significant, $F(1,34)=2.65$, $p=.113$. Planned comparisons revealed a main effect of sensibility for the right hemisphere, $F(1,34)=5.02$, $p<.04$, but not for the left hemisphere, $F<2$. No other interactions with the factor sensibility were found.

In sum, the analyses both for N400 mean amplitude and for N400 peak amplitude verified, first, that an N400 effect was present in Experiment 2, and, second, that the N400 sensibility effect in the reading task was restricted to the right hemisphere.

[4] Supplementary analyses were performed to assess the reliability of later ERP differences between conditions. To this aim we carried out analyses on the mean amplitude for the 600 to 800 ms epoch. The results of these analyses were as follows: Neither for the midline nor for the lateral sites, effects of sensibility (F values<1) or sensibility by site interactions (F values<1) were obtained. Furthermore, the lateral analyses did not yield other interactions of the factor sensibility with hemisphere, ROI, and/or site that could have pointed to reliable differences between conditions. From this, we conclude that no later differences were present between conditions.

3.2.3. Between-experiment analyses

The results of the individual experiments suggest that there are differences in the scalp distribution of the N400 sensibility effects between the judgment task and normal reading. To test for differences in the scalp distribution of the N400 sensibility effects, global analyses with experiment as between subject factor were performed in which the topographies of the sensibility effects were compared directly. These analyses were carried out on difference scores (mean amplitude for novel senseless critical words minus mean amplitude for novel sensible critical words in the 300 to 500 ms window) to remove the confounding effects of global differences in the size of the two kinds of N400 effects. The logic is that if there are differences between the topographies of the N400 sensibility effects between experiments, this should be signified by an interaction of the factor experiment with either site and/or hemisphere and ROI. The corresponding topographical maps showing the spatial scalp distribution of the N400 sensibility effects for Experiments 1 and 2 are presented in Fig. 4.

The between-experiment analyses for the midline sites did not yield an experiment by site interaction, $F<1$. This indicates that across the midline there were no reliable differences in the N400 sensibility patterns between experiments. The only effect that was found was a main effect of experiment, $F(1,63)=5.11$, $p<.05$, that reflected that overall mean amplitude was a bit more negative in Experiment 1 than in 2.

The between-experiment analysis for the lateral sites yielded main effects of experiment, $F(1,63)=4.31$, $p<.05$, and of hemisphere, $F(1,63)=4.16$, $p<.05$, reflecting small differences in overall amplitude between studies and hemispheres. More importantly, a two-way interaction between experiment and hemisphere was found, $F(1,63)=6.07$, $p<.05$. The interaction confirmed that in Experiment 2 an N400 sensibility effect was present for the right hemisphere but not for the left hemisphere (the mean N400 effect was −0.98 μV for the right hemisphere and −0.26 μV for the left hemisphere). For Experiment 1, there were no indications for differences in N400 sensibility effects between hemispheres, $F<1$ (the mean N400 effect was −0.423 μV and −0.427 μV for the left and the right hemisphere, respectively). The lateral analysis also yielded a ROI by site interaction, $F(4,252)=3.10$, $p<.04$. This interaction indicated that across experiments, N400 sensibility effects were largest (> than −0.8 μV) at bilateral posterior sites including the following electrodes LTP, RTP, P3, P4, P3p and P4p.

To sum up, the results of the between-experiment analyses for the lateral sites confirmed that the N400 sensibility effects in the judgment task and the reading task showed a different scalp distribution. Consistent with the topographical maps presented in Fig. 4, these analyses indicated that while the N400 sensibility effects in Experiment 1 were bilaterally distributed with maximal effects at centroparietal sites, the N400 sensibility effects in Experiment 2 showed a right hemisphere lateralization.

3.3. Discussion

The major finding of Experiment 2 was that novel sensible contexts that were matched in terms of familiarity to novel senseless contexts elicited an N400 effect. As in Experiment 1, the N400 sensibility effect resembled the standard N400 effect to associative and semantic relations in terms of the waveform and the timing of the effect (300 to 500 ms window). The N400 sensibility effect starts around 300 ms and peaks at

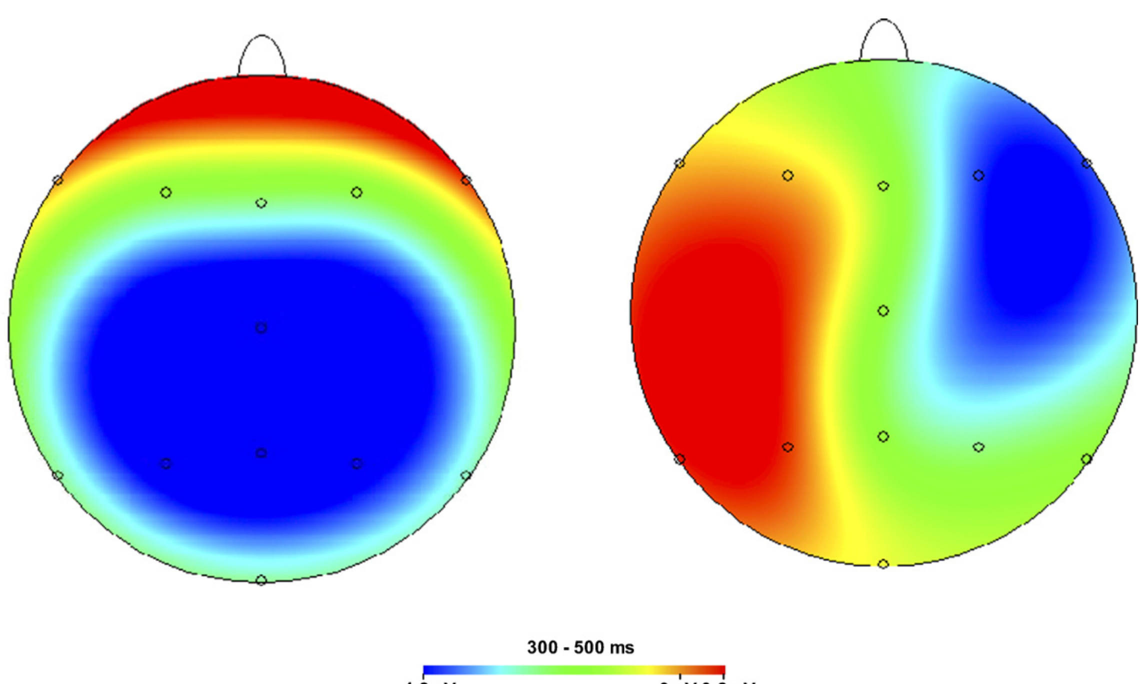

Fig. 4 – Topographical maps obtained by interpolation from 27 sites for the N400 effect (time window: 300–500 ms) for novel sensible meaning relations, separately for Experiment 1 (left side) and Experiment 2 (right side). Maps were computed from values resulting from the subtraction of related from unrelated waves.

400 ms. The main difference between studies was that the N400 sensibility effect in Experiment 2 was limited to the right hemisphere.

What are the implications of the N400 sensibility effects in the reading task employed in Experiment 2? While it is possible that artificial strategies induced by the sensibility judgment task, have contributed to the N400 effect in Experiment 1 this possibility can be ruled out in Experiment 2. In the latter study, participants did nothing else than what they do in daily life that is read for comprehension. The presence of an N400 effect in Experiment 2 that occurs in the same time frame as the standard N400 effect reveals that immediate integration of novel meanings forms part of the reading process. The ERP data of Experiment 2 provide further evidence for the embodied theory of language comprehension as presented by Glenberg and Robertson (1999, 2000). They support the view that the mental simulation process that plays a key role in embodied theories presents a competitor to abstract symbol theories of meaning. We will discuss the theoretical implications of this finding for current theories of meaning below.

Opposite to the N400 effect in the sensibility judgment task the effect during reading was restricted to the right hemisphere. The between-experiment analyses performed on the difference waves support this difference in the scalp distribution of the N400 effects between studies. It is important to point out that the standard N400 effect to associative and semantic relations typically is larger over the right than the left hemisphere. Nevertheless the standard N400 typically shows a bilateral distribution. What could the difference in scalp distribution between the N400 effects in Experiments 1 and 2 reflect? At face value, the lateralization of the N400 sensibility effect in Experiment 2 fits well with the idea that creative processes are mainly subserved by the right hemisphere. Such a specialization would be consistent with a growing number of studies that point to a special involvement of the right hemisphere in comprehending more complex language, such as discourse processing (St. George et al., 1999), comprehending metaphors (e.g., Bottini et al., 1994; Mashal et al., 2005) and getting jokes (Coulson and Wu, 2005; Goel and Dolan, 2001). The fact that in Experiment 1 the N400 effect was maximal at centroparietal midline sites may be explained in terms of strategies and/or decision-related processes induced by the artificial judgment task.

4. General discussion

The general goal of this article was to explore the nature of the process of new meaning creation. The more specific goal of the present experiments was to test abstract symbol theories against embodied theories of meaning. This was accomplished by comparing the online brain response to novel sensible contexts with that to novel senseless contexts that were matched in terms of familiarity/relatedness by Latent Semantic Analysis. The focus was on N400, which has been shown to be highly sensitive to semantic processing. The critical question for testing the two classes of theories against each other was whether novel sensible contexts relative to novel senseless contexts give rise to an N400 effect or not. The results on this point are clear: novel sensible contexts compared to novel senseless contexts elicit an N400 effect. Furthermore, this N400 effect generalizes from a task situation in which the sensibility of the critical materials was in the center of the attention (Experiment 1) to a normal reading situation (Experiment 2) indicating that the N400 sensibility effects reflect normal processes of language comprehension.

Our N400 results have implications for the following topics: current theories of meaning, the time course of the creation of new meanings, and the nature of the mental simulation process. We will now discuss these topics.

4.1. *Implications for current theories of meaning*

We have demonstrated that novel sensible meanings reliably elicit an N400 effect compared to novel senseless meanings. To our knowledge, this is the first report of an N400 effect for new meanings that have to be created on the spot. This is a theoretically important finding, because the N400 effect to novel sensible contexts cannot be explained by high-dimensional models of meanings (LSA of Landauer and Dumais (1997) and HAL of Burgess and Lund (2000)) or by other abstract symbol theories of meaning (e.g., Anderson et al., 1997; Collins and Loftus, 1975; Masson, 1995; Ratcliff and McKoon, 1988). According to abstract symbol theories, two concepts can only be meaningfully related if they are either associatively and/or semantically related. Given that the novel sensible concepts were neither associatively nor semantically related, these theories fail to explain why facilitation for these novel concepts relative to the control condition was observed. Current measures of associative strength or semantic similarity (i.e., free association, semantic category membership, local co-occurrence measures and global co-occurrence measures like LSA) therefore, cannot account for the N400 sensibility effect. Hence theories that employ these methods as measures of semantic memory, in particular, spreading activation models, compound cue models, distributed memory models and high-dimensional theories of meaning cannot explain the N400 effects to novel sensible contexts.

A fundamental problem that LSA and other abstract symbol theories of meaning face is how meaningless symbols come to take on meaning (Glenberg and Robertson, 2000; but see Burgess, 2000; Landauer and Dumais, 1997). As pointed out by Searle in his famous Chinese Room example (1980), manipulation of abstract amodal symbols produces more abstract symbols, not meaning. Glenberg and Robertson (2000) claim that they have (at least in part) solved this grounding problem in the embodied theory of action which they present as a competitor of abstract symbol theories. The starting point of the embodied approach of language comprehension is that language is grounded in perception and action. Meaning arises from our current physical interactions with the environment or previous encounters with objects in different kinds of environments. That is, present or past body by environment interactions guide us in how to think about, that is simulate the perceptual and action details required by a situation. Meaningfulness resides in our knowledge about the possibilities and limitations of the human body. Therefore, the finding of an N400 effect to novel sensible situations can be accounted for by the embodied theory of language, in particular by the Indexical Hypothesis of Glenberg and Robertson (1999, 2000).

This theory naturally explains why humans can make sense out of novel sensible situations but high-dimensional theories of meaning or abstract symbol theories cannot. The N400 data thus provide clear support for an embodied approach to meaning. They support the view that language comprehension is based on the physical interactions or imaginative simulations of interactions based on previous experiences that people have with their environment.

4.2. *The time course of the creation of new meanings*

ERPs provide a continuous record of the brain activity and are therefore an ideal method to track the time course of new meaning creation. The time course of the N400 sensibility effect relative to the standard N400 effect reported in the literature provides insight into the speed at which new meanings are constructed and integrated into a higher-order meaning representation of the current context. Most importantly, the N400 effect to novel meaning relations reported in Experiments 1 and 2 occurred in the same time frame (300–500 ms after critical word onset) as the typical N400 effect to associative and/or semantically related meaning relations. Previously, it has been shown that two kinds of world knowledge (well-known facts and scripts) are integrated as fast as associative and/or semantic meaning relations (Hagoort et al., 2004; Chwilla and Kolk, 2005; respectively). The present N400 results demonstrate that immediate integration is not restricted to familiar meaning relations that based on learning history are stored in long-term memory. Immediate integration also occurs for novel meaning relations, despite that the latter relations have never been experienced. As we will discuss below, we take the immediate N400 effects to novel sensible vs. novel senseless contexts to indicate that new meanings are not a product of processing strategies but are created on the spot.

Finally, the N400 sensibility effect in Experiment 2 during reading, while very similar in terms of the waveshape and timing to the standard N400 effect differed from the standard N400 effect in terms of the scalp distribution. The canonical N400 effect to semantic and associative relations is typically largest at centroparietal electrode sites and is slightly larger over right hemisphere than left hemisphere electrode sites (Kutas et al., 1988). The N400 sensibility effect in Experiment 2, on the other hand, was limited to the right hemisphere. This leads to the interesting hypothesis that the creation of new meanings that cannot be retrieved from long-term memory is realized by a somewhat different neuronal assemble than the integration of familiar meanings (see Fig. 4). However, it is clear that other brain-imaging methods like functional magnetic resonance imaging are needed to localize the brain areas involved in novel meaning creation.

4.3. *The nature of the mental simulation process*

The mental simulation process lies at the heart of embodied approaches to cognition in general, and language comprehension in specific. It could be argued that this mental simulation process represents some kind of strategy that is under attentional control and – opposite to symbolic processing which proceeds in a strict bottom up fashion and is automatic – is of a more controlled nature. Glenberg (2007) commented on the hypothesized nature of this mental simulation process. He argued that although simulations are some kind of imagery this does not necessarily imply that it concerns a consciously experienced imagery. He proposed further that this especially holds for familiar situations which may have resulted in simulations becoming automated. The present ERP data speak to this issue. First, the immediate N400 effect during reading shows that the mental simulation process is of a more automatic nature in that it is not subject to attentional control. Second, our N400 results reveal that affordances are immediately generated even for novel; that is, for unfamiliar situations that have never been experienced before. Apparently, deriving affordances – that is, simulating new actions like “to paddle with Frisbees” or “to drum with branches” occurs with the same ease than deriving affordances for familiar actions like to “paddle with paddles” or “to drum with sticks”. On the one hand, this ability seems astounding. On the other hand, if participants cannot help but search for coherence and meaning whenever encountering words (Chwilla et al., 2000) then the speed at which new meanings are established becomes understandable. Humans are experts in extracting meaning.

Note that work from developmental psychology provides evidence for the view that affordances are generated automatically. The Gibsons (1979, 1982, 1984) in their ecological theory of perception propose an *innate* mechanism for extracting important visual information from available stimulation. According to this nativist theory, any perceptually equipped animal can apprehend properties of the world that are relevant to action. They describe affordances as those action-related properties of the physical environment that afford, elicit, or suggest certain behavioral action. The organism’s needs, motivations and intentions in combination with the ecological context in which an element appears, indicate what an element affords. So, any organism will grasp what the environment affords in ecologically, valid, natural terrains. Consistent with this view Borghi (2005) proposed that at least for simple actions, visual input and object knowledge help subjects to extract affordances automatically. Recent studies in the field of language acquisition further bolster the claim that there exists a link between language and action. A nice example is a study of Smith (2005). In two experiments, she shows that action alters the shape categories formed by 2-year-old children. These data suggest a strong link between action on the one hand and categorization behavior on the other hand in very young children.

Taken together, the present N400 results, behavioral work from developmental psychology and language acquisition are all consistent with the claim that the hypothesized mental simulation process reflects a more automatic process.

4.4. *Conclusions*

From the ERP data, we draw three main conclusions: First, participants can establish new meaning relations that are not stored in semantic memory. Second, this is the first ERP study that shows that N400 is sensitive to such new meaning relations. Importantly, our N400 findings reveal that new meanings are created immediately – that is, in the same time frame as familiar meanings like associative and semantic relations. Third, the N400 sensibility effects reported in this article support embodied theories of meaning and challenge abstract symbol theories that can only discover meaningfulness by consulting stored symbolic knowledge. Whereas sensibility effects of familiarity can be explained on a purely symbolic basis, sensibility effects within completely unfamiliar contexts appear to force us to accept explanations in terms of embodied actions.

Table 2 – Matching of the critical and control items

	LSA (SSVs)[a], mean	Number of letters, mean	Number of syllables, mean	Mean frequency
Critical word				
Novel sensible	.119	7.92	2.48	23.64
Novel senseless	.138	7.62	2.34	26.82

[a] SSV refers to semantic similarity value.

5. Experimental procedures

5.1. *Experiment 1*

5.1.1. *Participants*

The participants were 30 students (age >18 and <30) at the Radboud University. All were native speakers of Dutch, had no reading disabilities, were right-handed, and had normal or corrected-to-normal vision. They were paid or received course credit for their participation.

5.1.2. *Materials*

Fifty-six Dutch scenarios were constructed. Each scenario consisted of one or more context-setting sentences followed by a test sentence (TS). A critical word in combination with a central concept rendered the TS given the context as either meaningful or not. Examples of scenarios translated from Dutch into American English are presented in Table 1. The critical word could occur in the midst of the sentence (in 16 sentences) or in sentence final position (in 34 sentences). The critical novel sensible words and novel senseless control words were matched on the following lexical characteristics: word length, number of syllables, and mean word frequency (see Table 2).

Two pretests were conducted to verify that the contexts were indeed novel. First, a paper and pencil test was carried out to verify that the novel sensible and novel senseless critical words were not produced in a cloze test. A separate group of 20 students participated in this study. The participants received a booklet that contained the paragraphs and the TSs. In the TSs, the critical words were omitted. Participant's task was to read the paragraph and the TS and write up as many as possible words that fit into the sentence given the context. Only those scenarios were accepted for which the critical novel sensible word or control word was never produced.

The second pretest consisted of the matching of the LSA semantic similarity values (SSVs) for the two sets of items. LSA is a technique that measures whether words co-occur in the same or in similar contexts. Although one could argue that LSA captures word associations rather than semantic relatedness per se, Chwilla and Kolk (2002) showed that LSA is a sensitive measure for detecting even subtle differences in (semantic) relatedness between words that were not associatively related. Following Glenberg and Robertson, we controlled that the SSVs for the novel sensible critical words were not larger than those for the control words. This was critical to rule out that novel sensible sentences were *more* familiar and/or semantically related than the control items. First, we computed the SSVs between a central concept in the sentence and the critical word that distinguished among the critical sentences (e.g., to paddle – Frisbees vs. to paddle – pullover). Second, we translated these items into American English and submitted them to LSA (Landauer and Dumais, 1997; Landauer et al., 1998) available on Internet at http://lsa.colorado.edu/. Specifically, we carried out a "Pairwise [term to term] comparison"[5] using the tasaALL space which corresponds to the first year college level. For this space, the LSA matrix is based on the occurrence of 92,409 unique terms in 37,651 contexts selected from "texts, novels, newspaper articles, and other information". The Single Value Decomposition (SVD) of this space yields vectors with a maximum of 419 dimensions. Only those sets of critical words and control words were selected for which the SSV was the same or larger for the control word than for the critical word. As Table 2 shows for the final set of 50 scenarios, the mean SSV for the control items was even larger than that for the critical items. Although

Table 1 – Examples of the stimuli

1. Context-setting sentences:	The boys found a canoe in the spare room. With this, they wanted to go canoeing on the canal whatever the costs. The fact that they could not find the paddles did not lead them to make up their mind. According to the boys, you do not at all need them.
Novel sensible test sentence:	They let the canoe into the water and paddled with *Frisbees*.
Novel senseless sentence:	They let the canoe into the water and paddled with *pullovers*.
2. Context-setting sentence:	The scouts wanted to make music at the campfire.
Novel sensible test sentence:	"The boys searched for *branches* with which they went *drumming* and had a lot of fun."
Novel senseless sentence:	"The boys searched for *bushes* with which they went *drumming* and had a lot of fun."

[5] Glenberg and Robertson (2000) also computed the LSA cosine for the whole sentence. Note that the cosines for a sentence are based on the average LSA vectors for the words comprising the sentence. Because the novel sensible vs. novel senseless sentences differed by only a word or two, the cosines were very similar and thus not very informative regarding the critical novelty condition. The other LSA measure, that we used (i.e., the cosine between a central concept and the concept that distinguishes among the critical sentences), however, revealed clear differences between the high cloze version of the sentences and both kinds of novel sentences. Most important for our present purposes is that the cosines for the two kinds of novel sentences did not differ.

the difference in SSVs was rather small it was significant, t_2 (49)=2.24, $p<.04$, indicating that the control items were a bit more familiar than the experimental items. Note that for our present purposes, this is not problematic because based on this difference in familiarity between items one would predict facilitation for the control items and not for the novel sensible items.

5.1.3. *Procedure*

Participants were seated in an electrically shielded chamber. Stimulus presentation and recording of the reaction time (RT) and error data was accomplished by an Apple Macintosh Power PC 7200/95. The stimuli were presented in upper-case black letters on a white background at moderate contrast (window of 8×2 cm; approximately 3.0°×0.8° of arc). The test sentence was presented in serial visual presentation mode at the center of a PC monitor. Word duration was 345 ms and the stimulus-onset asynchrony (SOA) was 645 ms. Sentence final words were followed by a full stop.

The participant got a booklet with the scenarios numbered from 1 through 50 and sensibility rating scales printed on separate pages. This booklet was placed in front of the participant on a small table such that (s)he could easily turn the pages without making many movements. The sequence of a single trial was as follows: The first set of 5 items consisted of practice trials. The experiment started when the text "Read practice paragraph 1" occurred on the screen. The subject read the paragraph. When ready (s)he pushed a button which triggered the presentation of a fixation cross for 500 ms. The cross was followed by the TS which was presented at the center of a monitor in RSVP mode. The participant was instructed to avoid making eye-movements after the fixation cross up to about 1.5 s after the sentence final word. After that the participant had to turn the page on which the paragraph was printed and indicated on a scale from 1 through 7 how sensible the TS was (7 indicates very meaningful and 1 very senseless) given the context. When ready (s)he looked at the screen where the message "Read practice paragraph 2" was presented. Then (s)he completed the series of steps described above until the last practice trial was completed. There was a brief pause in which the participant could ask final questions. After that the experiment was started.

5.1.4. *Reaction time pilot study*

A RT pilot study was conducted to investigate if novel sensible contexts are processed faster and/or more accurately than novel senseless contexts. For the RT study, a response device containing two push buttons was fixed on a small table in front of the participant. The instruction was the same as that for the ERP study, except that participants ($N=24$) had to indicate as fast as possible after reading the sentence final word whether the TS was meaningful (right-hand button press) or not (left-hand button press). The main result was that there were no reliable differences in reaction times ($t=1.64$, $p=.108$; mean RT for novel sensible and senseless TSs was 2110 ms [SD=634] and 1893 ms [SD=858], respectively) or error rates ($t<1$; 11% and 9% for the sensible and senseless TSs, respectively) between novel sensible and novel senseless TSs. Note that overall participants tended to be *slower* and less accurate to novel sensible than to the control items.

5.1.5. *Electrophysiological recording*

The electroencephalogram (EEG) was recorded with 27 tin electrodes mounted in an elastic cap (Electro-cap International, Eaton, OH) from 5 midline and 22 lateral sites (see Fig. 1 for the montage). All EEG channels were referenced to the left mastoid, and the right mastoid was employed as an active recording channel. Before the analyses, the signals were re-referenced to the average of the right and the left mastoids. Vertical EOG was measured by placing an electrode above and below the right eye and the horizontal EOG was recorded via a right to left canthal montage. Electrode impedance was less than 3 kΩ. The signals were amplified (time constant=8 s, bandpass=.02–30 Hz), and digitized on-line at 200 Hz.

Analyses of variance (ANOVAs) were performed with repeated measures on sensibility (sensible vs. senseless). Before analysis, EEG and EOG records were examined for artifacts and for excessive EOG amplitude (>100 μV) from 100 ms preceding the critical word to 1 s after its onset. Only artifact-free trials were included in the average. ERPs were averaged time-locked to critical word onset, relative to a 100 ms pre-word baseline. Mean N400 amplitude was measured in the 300 to 500 ms window following critical word onset which corresponds to the standard N400 time window. Supplementary N400 analyses using a smaller N400 window, in particular the 400 to 500 ms epoch, were also performed because in a previous study on integration of script knowledge we observed maximal N400 effects in this epoch. The results of these analyses will only be reported if they provide additional information.

Because one may expect a larger variability in the N400 effects, supplementary analyses on N400 peak amplitude, measured in the 300 to 500 ms window from critical word onset, were performed. For both measures (mean amplitude and peak amplitude) two kinds of analyses were performed. For the midline sites analyses were carried out with sensibility (2 levels), and site (5 levels) as factors. For the lateral sites, analyses were conducted with sensibility as factor using a region of interest (ROI: anterior vs. posterior) by hemisphere by lateral site (F7a/F3a/F7/F3/LAT vs. LTP/P3/P3p/T5/OL vs. F8a/F4a/F8/F4/RAT vs. RTP/P4/P4p/T6/OR) design. The electrode montage is displayed in Fig. 1.

In the literature, it has been claimed that the right hemisphere plays a special role in creative behavior. To test for possible hemispheric differences, separate analyses were performed for the left and the right hemisphere. In all analyses, the estimated Greenhouse and Geisser coefficient epsilon was used to correct for violations of the assumption of sphericity. All reported p-values are based on corrected degrees of freedom but, to aid the reader in interpreting our statistical design, the stated degrees of freedom are uncorrected.

5.2. Experiment 2

5.2.1. *Participants*

Thirty-five students who fulfilled the same criteria as those in Experiment 1 participated.

5.2.2. *Apparatus and materials*

The apparatus, presentation conditions and materials were the same as in Experiment 1.

5.2.3. *Procedure*

The procedure was the same as in Experiment 1 except that subjects did nothing else than read the sentences.

5.2.4. *Data analysis*

Other aspects of the analyses were the same as in Experiment 1.

Acknowledgments

The research was supported by grant 560-256-085 from the Netherlands Organization for Scientific Research (NWO). Portions of this research were presented at the Annual Meetings of Cognitive Neuroscience in San Francisco, 2004 and New York, 2005. We thank two anonymous reviewers for helpful comments on previous versions of this article. We are grateful to Arthur Glenberg for providing us his materials. Nine of our 50 Dutch scenarios were translations from the original Glenberg and Robertson paper (2000). We want to thank the ERG group of NICI, in particular Pascal de Water for technical assistance and Inge Leunissen for running additional subjects. Last but not least, we thank Uli Chwilla for preparing all figures. This article is dedicated to the *art of haiku*, a brief expression of transient enlightenment in which we see the living reality of objects.

REFERENCES

Anderson, J.R., Matessa, M., Lebiere, C., 1997. ACT-R: A theory of higher level cognition and its relation to visual attention. Hum.–Comput. Interact. Spec. Issue: Cogn. Arch. Hum.–Comput. Interact. 12 (4), 439–462.

Barsalou, L.W., 1993. Flexibility, structure, and linguistic vagary in concepts: manifestations of a compositional system of perceptual symbols. In: Collins, A.C., Gathercole, S.E., Conway, M.A. (Eds.), Theories of Memories. Erlbaum.

Barsalou, L.W., 1999. Perceptual symbol systems. Behav. Brain Sci. 22, 577–609.

Bodner, G.E., Masson, M.E.J., 2003. Beyond spreading activation: an influence of relatedness proportion on masked semantic priming. Psychon. Bull. Rev. 10, 645–652.

Borghi, A.M., 2005. Object concepts and action. In: Pecher, D., Zwaan, R.A. (Eds.), Grounding Cognition. The Role of Perception and Action in Memory, Language, and Thinking. Cambridge University Press.

Bottini, G., Corcoran, R., Sterzi, R., Paulesu, E., Schenone, P., Scarpa, P., Frackowiak, R.S.J., Frith, C.D., 1994. The role of the right-hemisphere in the interpretation of figurative aspects of language – a positron emission tomography activation study. Brain 117, 1241–1253.

Burgess, C., 2000. Theory and operational definitions in computational memory models: a response to Glenberg and Robertson. J. Mem. Lang. 43, 402–408.

Burgess, C., Lund, K., 2000. The dynamics of meaning in memory. In: Dietrich, E., Arthur, B. (Eds.), Cognitive Dynamics: Conceptual and Representational Change in Humans and Machines. Erlbaum, Mahwah, NJ, pp. 117–156.

Chwilla, D.J., Kolk, H.H.J., 2002. Three-step priming in lexical decision. Mem. Cogn. 30, 217–225.

Chwilla, D.J., Kolk, H.H.J., 2005. Accessing world knowledge: evidence from N400 and reaction time priming. Cogn. Brain Res. 25, 589–606.

Chwilla, D.J., Hagoort, P., Brown, C.M., 1998. The mechanism underlying backward priming in a lexical decision task: spreading activation versus semantic matching. Q. J. Exp. Psychol., A 51, 531–560.

Chwilla, D.J., Kolk, H.H.J., Mulder, G., 2000. Mediated priming in the lexical decision task: evidence from event-related potentials and reaction time. J. Mem. Lang. 42, 314–341.

Collins, A.M., Loftus, E.F., 1975. A spreading-activation theory of semantic processing. Psychol. Rev. 82, 407–428.

Coulson, S., Wu, Y.C., 2005. Right hemisphere activation of joke-related information: an event-related brain potential study. J. Cogn. Neurosci. 17, 494–506.

Fodor, J., 1983. The Modularity of Mind. MIT Press, Cambridge, MA.

Friederici, A.D., 1995. The time-course of syntactic activation during language processing: a model based on neuropsychological and neurophysiological data. Brain Lang. 50, 259–281.

Gibbs Jr., R.W., 2005. Embodiment in metaphorical imagination. In: Pecher, D., Zwaan, R.A. (Eds.), Grounding Cognition. The Role of Perception and Action in Memory, Language, and Thinking. Cambridge University Press.

Gibson, J.J., 1979. The Ecological Approach to Visual Perception. Houghton Mifflin, Boston.

Gibson, E.J., 1982. The concept of affordances in development: the renascence of functionalism. In: Collins, W.A. (Ed.), Minnesota Symposia on Child Psychology: Vol. 15. The Concept of Development. Erlbaum, Hillsdale, NJ.

Gibson, E.J., 1984. Perceptual development from the ecological approach. In: Lamb, M.E., Brown, A.L., Rogaff, B. (Eds.), Advances in Developmental Psychology. Erlbaum, Hillsdale, NJ.

Glenberg, A.M., 1997. What memory is for. Behav. Brain Sci. 20, 1–55.

Glenberg, A.M., Robertson, D.A., 1999. Indexical understanding of instructions. Discourse Processes 28, 1–26.

Glenberg, A.M., Robertson, D.A., 2000. Symbol grounding and meaning: a comparison of high-dimensional and embodied theories of meaning. J. Mem. Lang. 43, 379–401.

Glenberg, A.M., Kaschak, M.P., 2002. Grounding language in action. Psychon. Bull. Rev. 9, 558–565.

Glenberg, A.M., Havas, D., Becker, R., Rinck, M., 2005. In: Pecher, D., Zwaan, R.A. (Eds.), Grounding Cognition. The Role of Perception and Action in Memory, Language, and Thinking. Cambridge University Press.

Glenberg, A.M., 2007. Language and action: creating sensible combinations of ideas. In: Gaskell, M.G. (Ed.), The Oxford Handbook of Psycholinguistics. Oxford University Press, Oxford, pp. 361–370.

Goel, V., Dolan, R.J., 2001. The functional anatomy of humor: segregating cognitive and affective components. Nature Neuroscience 4, 237–238.

Hagoort, P., Hald, L., Petersson, M., 2004. Integration of word meaning and world knowledge in language comprehension. Science 304, 438–441.

Harnad, S., 1990. The symbol grounding problem. Physica, D 42, 335–346.

Hess, D.J., Foss, D.J., Carroll, P., 1995. Effects of global and local context on lexical processing during language comprehension. J. Exp. Psychol. Gen. 124, 62–82.

Holcomb, P.J., 1993. Semantic priming and stimulus degradation: implications for the role of the N400 in language processing. Psychophysiology 30, 47–61.

Hutchison, K., 2003. Is semantic priming due to association strength or featural overlap? A "micro-analytic" review. Psychon. Bull. Rev. 10, 785–813.

Jackendoff, R., 1987. Consciousness and the Computational Mind. MIT Press, Cambridge, MA.

Katz, J.J., Fodor, J.A., 1963. The structure of a semantic theory. Language 39, 170–210.

Kiefer, M., 2002. The N400 is modulated by unconsciously perceived masked words: further evidence for an automatic

spreading activation account of N400 priming effects. Cogn. Brain Res. 13, 27–39.
Kintsch, W., 1988. The role of knowledge in discourse comprehension. A construction–integration model. Psychol. Rev. 95, 163–182.
Kutas, M., Van Petten, C., Besson, M., 1988. Event-related potential asymmetries during the reading of sentences Electroencephalogr. Clin. Neurophysiol. 69, 218–233.
Lakoff, G., 1987. Women, Fire and Dangerous Things: What Categories Reveal about the Mind. University of Chicago Press, Chicago.
Landauer, T.K., Dumais, S.T., 1997. A solution to Plato's problem: the latent semantic analysis theory of acquisition, induction, and representations of knowledge. Psychol. Rev. 104, 211–240.
Landauer, T.K., Foltz, P.W., Laham, D., 1998. Introduction to latent semantic analysis. Discourse Processes 25, 259–284.
Mashal, N., Faust, M., Hendler, T., 2005. The role of the right hemisphere in processing nonsalient metaphorical meanings: application of principal components analysis to fMRI data. Neuropsychologia 43, 2084–2100.
Masson, M.E.J., 1995. A distributed memory model of semantic priming. J. Exper. Psychol., Learn., Mem., Cogn. 21, 3–23.
McNamara, T.P., 2005. Semantic Priming: Perspectives from Memory and Word Recognition. Essays in Cognitive Psychology. Psychology Press New York and Hove.
Pecher, D., Zwaan, R.A., 2005. Grounding Cognition. The Role of Perception and Action in Memory, Language, and Thinking. Cambridge University Press.
Ratcliff, R., McKoon, G., 1988. A retrieval theory of priming in memory. Psychol. Rev. 95, 385–408.
Rolke, B., Heil, M., Streb, J., Hennighausen, E., 2001. Missed prime words within the attentional blink evoke an N400 semantic priming effect. Psychophysiology 38, 165–174.
Schank, R.C., Abelson, R.P., 1977. Scripts, Plans, Goals and Understanding. Erlbaum, Hillsdale, NJ.
Searle, J.R., 1980. Minds, brains, and programs. Behav. Brain Sci. 3, 417–457.
Smith, L.B., 2005. Action alters shape categories. Cogn. Sci. 29, 665–679.
Stanfield, R.A., Zwaan, R.A., 2001. The effect of implied orientation derived from verbal context on picture recognition. Psychol. Sci. 12, 153–156.
Stenberg, G., Lindgren, M., Johansson, M., Olsson, A., Rosen, I., 2000. Semantic processing without conscious identification: Evidence from event-related potentials. J. Exper. Psychol. Learn., Mem., Cogn. 26, 973–1004.
St. George, M., et al., 1999. Semantic integration in reading: engagement of the right hemisphere during discourse processing. Brain 122, 1317–1325.
Van Berkum, J.J.A., Brown, C.M., Hagoort, P., 1999. Early referential context effects in sentence processing: evidence from event-related brain potentials. J. Mem. Lang. 41, 147–182.
Van Herten, M., Chwilla, D.J., Kolk, H.H.J., 2006. When heuristics clash with parsing routines: ERP evidence for conflict monitoring in sentence perception. J. Cogn. Neurosci. 18, 1181–1197.
Zwaan, R.A., Madden, C.J., 2005. Embodied sentence comprehension. In: Pecher, D., Zwaan, R.A. (Eds.), Grounding Cognition. The Role of Perception and Action in Memory, Language, and Thinking. Cambridge University Press.
Zwaan, R.A., Radvansky, G.A., 1998. Situation models in language comprehension and memory. Psychol. Bull. 123, 162–185.
Yeh, W., Barsalou, L.W., 2006. The situated nature of concepts. Am. J. Psychol. 119, 349–384.