Semantic Illusion in Sentence Processing: a Right-Hemisphere Mechanism?

John C. J. Hoeks (hoeks@let.rug.nl)

Ingeborg Prinsen (prinsen@let.rug.nl)

Laurie A. Stowe (stowe@let.rug.nl)

Abstract

Recent research on sentence processing using ERPs (Event Related brain potentials) has shown that there are situations in which the semantic relationships between words in a sentence are so strong that they can block the semantic interpretation that is actually prescribed by the syntactic structure of that sentence (Hoeks, Stowe, & Doedens, 2003; Kolk, Chwilla, van Herten, & Oor, in press). As syntactic processing is the assumed province of the left hemisphere (LH), it was hypothesized that this so-called ‘semantic illusion’ might result from a transient but apparently rather influential non-syntactic sentence representation formed in the right hemisphere (RH). Two reaction time experiments using the Divided Visual Field paradigm only partially supported this hypothesis, as they showed that it is the LH that is most sensitive to semantic illusion.

Introduction

Readers do not wait with the interpretation of a sentence until they have received the final word. On the contrary, the process of understanding sentences occurs in a highly incremental fashion, approximately as each word is encountered (e.g., Altmann & Steedman, 1988). The partial sentence representations that are produced by this continuous process of interpretation have been shown to facilitate the processing of upcoming words. For instance, Duffy, Henderson, & Morris (1989) showed with eye tracking that if a new word is semantically related to two or more words in the preceding sentence context, it is fixated for a significantly shorter time than in a semantically ‘neutral’ context. This kind of facilitated processing, possibly originating from some kind of rather ‘coarse’ semantic representation based on all preceding lexical items taken together, was called ‘lexical’ facilitation.

Following up on Duffy et al., Morris (1994) showed that semantic relations indeed play an important facilitatory role in sentence processing, but that this effect is mediated by the ‘message-level’ representation of a given sentence, that is, the representation in which both semantic and syntactic information are taken into account. To illustrate, in a sentence such as The gardener talked as the barber trimmed the MUSTACHE, the semantically related words remained the same and in approximately the same position, but the representation at the sentence-level changed considerably before the final word is reached: Not the barber, but the gardener is doing the trimming here. Morris showed that mustache is not facilitated under these circumstances. In other words, facilitation is governed by the message-level representation of a sentence.

However, a recent study by Hoeks, Stowe, & Doedens (2003) using ERPs (Event Related brain Potentials) found evidence that seems to contradict this ‘message-level hypothesis’. In their experiment they used the N400 amplitude evoked by the final word was used as a dependent measure; the N400 being a negative component of the ERP that peaks some 400 ms after presentation of a stimulus and is highly sensitive to semantic processing: the easier it is to process a given item semantically, the smaller the N400 (e.g., Kutas & Hillyard, 1984). Hoeks et al. used materials like the following (lit. = literal English translation of the Dutch example sentence):

1. Plausible & Related
Het brood werd door de bakkers GEBAKKEN.
lit. The bread was by the bakers BAKED.

2. Implausible & Related
Het brood heeft de bakkers GEBAKKEN.
lit. The bread has the bakers BAKED.

3. Implausible & Unrelated
Het brood heeft de bakkers BEDREIGD.
lit. The bread has the bakers THREATENED.

According to the ‘message-level hypothesis’, N400 amplitude to the target word baked in sentence 1 should be smallest, as this word is very easy to process; in contrast, the final words in sentences 2 and 3 should give rise to much larger amplitudes because they obviously do not fit into the existing message-level representation; both are equally implausible as ascertained in a separate rating study. Surprisingly, however, no significant difference in N400 amplitude was found for the final words of sentences 1 and 2, while both differed significantly from the N400 elicited by control sentence 3. It was only some 700 ms after presentation of the final word that the ERP waveforms
for plausible sentence 1 and implausible sentence 2 started to diverge (a positive shift was observed that might indicate processing difficulty related to, e.g., syntactic structure building, or reanalysis). Thus it is very likely that, at least temporarily, sentence 2 was wrongly taken as highly plausible. In other words, these results point to a phenomenon that may be called a temporary ‘semantic illusion’. See Figure 1 for the results (in microvolts) at electrode Pz (i.e., an electrode near the top of the head that is generally highly sensitive to modulations of the N400).

![Figure 1. ERP waveforms from the Hoeks et al. study](image)

Note that after reading these sentences participants were required to make a plausibility judgment. The majority of semantic illusion sentences were correctly classified as being implausible (i.e., 89%), indicating that the illusion is really a temporary phenomenon.

Hoeks et al. argued that in order to find a semantic illusion effect, two conditions should be met. First, there must be some problem in the timely construction of a message-level representation. For sentence 2 this difficulty might arise from the fact that the thematic relations (i.e., ‘who is doing what to whom’) in this sentence are not at all clear: the syntax prescribes that the inanimate entity (the bread) should do something to the animate entities (the bakers), which is not the usual state of affairs. The second condition is that all words in the sentence should fit together semantically (not necessarily associatively; more like fitting into one concept or scenario) thus facilitating the processing of the target word. So if the construction of a valid message-level representation is hampered or seriously delayed, there can be significant lexical facilitation if the words fit together.

This was not the only demonstration of the semantic illusion effect. In fact, the Hoeks et al. (2003) results were practically replicated by the results of another recent study by Kolk and co-workers (Kolk, Chwilla, van Herten, & Oor, in press). They used sentences such as the following, and measured the N400 on the target word *hunted*:

4. De stroper die op de vossen joeg ...
   lit. The poacher that on the foxes hunted ...

5. De vos die op de stropers joeg ...
   lit. The fox that on the poachers hunted ...

As in the Hoeks et al. study, Kolk et al. did not find any evidence for an N400 difference between these sentences, even though sentences such as 5 were rated as highly implausible. Instead, they found a late positive component they interpreted as indicative of syntactic processing difficulty. Again, the prerequisites for the semantic illusion effect are present: thematic processing difficulty (cf., foxes that hunt poachers in 5 vs. poachers that hunt foxes in 4) and words that are highly semantically related.

To briefly summarize, we have seen that during some phase of sentence interpretation, effects of semantic relatedness can ‘overrule’ syntactic structure, that is, even if the syntactic structure of a sentence permits only one interpretation, a strong semantic relatedness between the content words in the sentence can temporarily overturn or even block this obligatory interpretation. Perhaps it is possible that there are actually two mechanisms of sentence interpretation: one responsible for a message-level representation (with syntax) and one for a coarse-grained semantic representation that does not need syntax. And perhaps each of these mechanisms is housed in a separate hemisphere. As syntactic processing is assumed to be the province of the left hemisphere (LH), it may be hypothesized that the 'semantic illusion' results from a transient but apparently influential coarse-grained semantic sentence representation formed in the right hemisphere (RH). This might seem farfetched, but in the next paragraph we will explain why we think this is a plausible hypothesis.

A Right Hemisphere Phenomenon?

In the last thirty years a lot of research has been dedicated to unraveling the linguistic capabilities of the RH, as compared to the language-dominant LH. The general picture that emerges from the literature is that though the two hemispheres collaborate closely during language processing, they have a specific division of labour. For instance, it has been argued that left hemisphere language processing takes place at the message-level where both semantic and syntactic information are integrated, while processing in the RH proceeds in a more global manner, and is more geared toward semantic coherence (e.g., Beeman et al., 1994).

Given the evidence for the different modes of language processing in the two hemispheres, Hoeks et al. (2003) speculated that there might actually be two mechanisms for sentence interpretation that, in spite of their close cooperation, are nevertheless dedicated to different aspects of the interpretation process. One of these mechanisms, in the LH, would then be responsible for creating the message-level representation (i.e., with the use of syntax), whereas the other, located in the RH, continuously creates a coarse-grained semantic representation into which all content words are integrated, thus more or less representing the ‘gist’ of the sentence. If, for some reason or other, the LH is not able to produce a valid message-level representation quickly enough, the RH temporarily takes over to guide the integration of incoming lexical items. This could explain why there is a semantic illusion effect in the experiments discussed above.

To test the hypothesis that a RH mechanism is responsible for the semantic illusion effect, two experiments were performed using the divided visual
field technique. In these experiments, the final word of a centrally presented sentence was flashed in either the left visual field (LVF, directly connected to the RH) or the right visual field (RVF, directly connected to the LH). Information presented in either of the two hemifields is initially available only to the contralateral hemisphere. In Experiment 1 participants were asked to judge whether the final word was a real Dutch word or not (lexical decision); in Experiment 2 they judged whether or not a sentence made sense (plausibility judgment).

**Experiment 1**

In the first experiment the lexical decision technique was used where participants are presented with letterstrings and asked to decide whether a specific letterstring is an existing word in their language (here: Dutch) or not. This task has been shown to be sensitive to the semantic processing of a word in relation to its preceding context in the following sense: If a target word is semantically congruent with the preceding context, deciding that it is a word occurs faster than in a neutral context; on the other hand, if the word does not fit into the context, decision times are longer as compared to a neutral context (e.g., O’Seaghdha, 1997). So in general target words occurring in implausible sentences receive much longer decision times than when occurring in plausible sentences. In this experiment we will test whether this is true for all implausible sentences or whether we can find evidence for the semantic illusion effect. In that case we expect shorter decision times and less errors in the implausible-related condition than in the implausible control condition, possibly even close to the performance in the plausible-related condition. If there is a semantic illusion effect, we expect it to be predominantly present in the LVF/RH.

**Method**

**Participants**  Fourty-eight native speakers of Dutch were paid for participating in this experiment (36 female; mean age 21 years, age range 18-30). All were currently receiving a university education. Participants were righthanded, had normal or corrected-to-normal vision and did not report having (had) neurological problems.

**Materials & Design**  The complete set of materials of the Hoeks et al. (2003) study were used, consisting of 112 sentences. In the present study there were three factors. The first factor, sentence type, had three levels: 1) plausible with related final word, 2) implausible with related final word (the ‘semantic illusion’ condition), and 3) implausible with unrelated final word. Examples of these three conditions are given by sentences 1-3 respectively (see Introduction). For the present experiment two other factors were added: field, with levels LVF and RVF, and word, with levels word and pseudoword. For the pseudoword condition, pseudowords were constructed that could have been words according to Dutch phonotactic constraints but that were not (e.g., "Het brood werd door de bakkers geperf"").

Experimental lists were constructed with an equal number of items per condition, and no list containing more than one version of a given item. Filler sentences were added such that each list contained an equal number of plausible and implausible items. The order in which experimental and filler items appeared was determined semi-randomly and was the same for each list. Each list was presented to an equal number of participants and each participant only saw one list.

All target stimuli (i.e., final words of the experimental sentences) were carefully matched with respect to length (in characters), frequency of occurrence, and plausibility of the sentence was taken into account (rated on a scale from 1=highly implausible, to 5=highly plausible).

<table>
<thead>
<tr>
<th></th>
<th>length</th>
<th>logfreq</th>
<th>implaus</th>
<th>plaus</th>
</tr>
</thead>
<tbody>
<tr>
<td>related</td>
<td>8.1 (1.5)</td>
<td>1.4 (0.7)</td>
<td>1.4 (0.4)</td>
<td>4.5 (0.5)</td>
</tr>
<tr>
<td>unrelated</td>
<td>8.0 (1.2)</td>
<td>1.4 (0.7)</td>
<td>1.4 (0.5)</td>
<td>---</td>
</tr>
</tbody>
</table>

Note. implaus = plausibility rating for related and unrelated words in plausible sentences; plaus = plausibility rating for related words in plausible sentences.

A practice session consisting of 47 items preceded the actual experiment.

**Procedure**  Participants were seated behind a computer screen (distance to screen = 60 cm) in a sound-proof cabin. A chin-rest helped them to keep their head in the same optimal viewing position while doing the experiment. Sentences were presented word-by-word (480 ms per word; 246 ms on screen, followed by a 234 ms blank screen) in the center of the screen, except for the final word which was flashed either left or right of the center (at a horizontal visual angle of 2°) for a duration of 200 ms. Participants were instructed to press either a red ("word") or a green ("non-word") button on a keyboard with their right index finger or their right middle finger, respectively.

Participants were asked to read the sentences carefully and to respond as quickly as possible without compromising accuracy. After a response was made, the word *knipper* (i.e., "blink") appeared for 3 sec., as participants were encouraged not to blink during reading to avoid them missing any of the final words. After that, an asterisk appeared for 1 sec. to indicate that the next sentence was imminent.

A video-camera was installed on top of the computer screen to keep track of the participants' eye movements. Participants were instructed not to move their eyes to the words flashed in either LVF or RVF, and were trained during the practice session until they no longer did so. During the actual experiment the participants' eye movements were continuously monitored and saccades were counted. No participant was observed to make more than 10 saccades during the entire experiment (i.e., including filler sentences).
Sentences were presented in four blocks of 56 sentences each. After each block participants could take a short break. In all, the experiment took approximately 45 min.

Results

Analysis The reaction time data were screened for outliers. In addition, mean error percentages in each condition were calculated for both participants and items. Figure 2 presents mean RTs and error percentages per condition of Experiment 1. Two ANOVAs (Analyses of Variance) were performed. The first analysis involved testing whether the semantic illusion effect, that is, shorter RTs and less errors for implausible-related sentences than for implausible-unrelated sentences was different for RH than for LH; this is the field (2) x relatedness (2) analysis. The second analysis looked into the difference between the plausible condition and the two implausible conditions; this is the field (2) x sentence type (3) analysis. These two analyses were performed on reaction times and error percentages. Normally these analyses are based on both participant (F1) and item means (F2). However, due to the relatively large number of errors made by the participants there were not enough observations per condition per item to perform statistically valid analyses on the item means for the reaction times. Therefore we will only report the outcomes of the analyses by participants in the paragraph on reaction times.

![Figure 2. Reaction times (upper panel) and Error percentages (lower panel) for Experiment 1.](image)

P-R=plausible-related; I-R=implausible-related; I-U=implausible-unrelated. LVF=left visual field; RVF=right visual field.

Reaction Times The field/relatedness analysis produced an interaction between field and relatedness on reaction times, which was marginally significant (F(1,42)=2.9; p=.096). A post-hoc analysis revealed a significant semantic illusion effect for stimuli presented to the RVF (F(1,44)=9.8; p<.005), and a statistical null-effect for the LVF (F<1). As to the other effects, there was a marginally significant main effect of relatedness (F(1,42)=3.2; p=.08), reflecting a trend for shorter RTs in the implausible-related condition than in the implausible-unrelated condition (a difference of 43 ms). The main effect of field, finally, was highly significant (F(1,42)=9.6; p<.005): RTs were longer when target words were presented to the LVF (1166 ms) than to the RVF (1086 ms), a difference of 80 ms.

No interaction was found in the field/sentence type analysis (F < 1). There was a main effect of field (F(1,42)=16.2; p<.001) and of sentence type (F(1,42)=27.5; p<.001). The first main effect indicated the slower processing of LVF stimuli (LVF: 1120 ms; RVF: 1030 ms), the main effect of sentence type reflected the much faster times for the plausible sentences (implausible-related: 1104 ms; implausible-unrelated: 1147 ms; plausible-related: 974 ms).

Error Percentages The field/relatedness analysis did not reveal significant interactions between field and relatedness in the error data (F1(1,47)=1.8; p=.19; F2<1). However, both main effects, of field (F1(1,47)=30.5; p<.001; F2(1,111)=22.1; p<.001) and relatedness were highly significant (F1(1,47)=33.8; p<.001; F2(1,111)=44.3; p<.001): target words presented to LVF were more often incorrectly classified as pseudo-words (i.e., 33 % of the time) than when presented to RVF (22 %). More importantly, target words that were presented in the context of an implausible-related sentence were less often seen as pseudo-words (20 %) than when occurring in the implausible controls with unrelated words (35 %). In other words, the semantic illusion effect in terms of a reduction in percentage of errors during lexical decision turned out to be approximately equally large in LH and RH.

The field/sentence type analysis did not produce significant interactions of field and sentence type (both Fs <1). There was a main effect of field (F(1,47)=43.9; p<.001; F2(1,111)=36.4; p<.001) reflecting more errors for LVF (27 %) than for RVF (16 %), and there was also a significant main effect of sentence type (F(1,47)=9.3; p<.005; F2(1,111)=36.4; p<.001) showing most errors for the implausible control sentences (35 %), considerably less errors for the semantic illusion sentences (20 %), and fewest errors for the plausible sentences (9 %).

Discussion

This experiment yielded two important results. First, in both reaction times and error percentages evidence was found for the semantic illusion effect: decision times were shorter and error percentages were smaller for the implausible-related sentences than for the implausible controls. Secondly, contrary to our hypothesis, the semantic illusion effect did not appear to be an exclusive RH phenomenon at all, as it was found in both hemispheres in terms of error percentages, and only in the LH when reaction times were concerned.

Looking only at the error data we might conclude that there is not much difference in processing between the
two hemispheres, as both show a large semantic illusion effect. However, if we consider the reaction time data, the strict lateralization of the semantic illusion effect suggests that the hemispheres do process language differently, and that it is the LH that is most sensitive to semantic information, not the RH.

This dissociation between the two dependent measures certainly suggests that each of these measures must reflect a different aspect of processing. It is possible that reaction time is a more on-line measure of sentence processing, while the errors might reflect some later stage, e.g., the process of decision making or error monitoring, which is also a very important aspect of the lexical decision task that was used (cf. Kaplan & Zaidel, 2001). To determine the extent to which our present pattern of results is specific to the particular task that was used in the experiment, we decided to try and replicate the present findings with a different, and probably less complicated task: plausibility judgment.

Experiment 2

This experiment used plausibility judgment time and judgment errors as its dependent measures. The logic behind the experiment is simple: the semantic illusion effect will tempt readers to say that implausible sentences with related words are actually plausible, which should lead to longer judgment times and to more errors than for the control sentences which are equally implausible (e.g., sentence 3). Furthermore, if the semantic illusion is the consequence of an integrative mechanism in the RH, and the results from the previous experiment are highly task-dependent, this pattern of longer reaction times and more errors is expected to obtain when the stimulus is presented in the LVF/RH, but not, or substantially less, when presented in the RVF/LH.

Method

Participants Forty native speakers of Dutch were paid for participating in this experiment (32 female; mean age 21 years, age range 18-26); all were currently receiving a university education. Participants were righthanded, had normal or corrected-to-normal vision and did not report having (had) neurological problems.

Materials, Design, & Procedure Materials and design were the same as in Experiment 1, now without the pseudowords. Procedure, and other experimental details were also the same as in Experiment 1, except that participants had to decide whether the sentence was plausible or not.

Results

Analysis After screening for outliers (see Experiment 1), mean RTs and mean error percentages in each condition were calculated for both participants and items (see Figure 3). Both the field (2) x relatedness (2) and the field (2) x sentence type (3) ANOVA were performed (see Experiment 1)

Reaction Times The field/relatedness analysis on reaction times did not produce the predicted interaction of field and semantic relatedness (Fs < 1). Although there was no main effect of relatedness, the main effect of field was highly significant, as RTs were much longer when target words were presented to the LVF (1238 ms) than to the RVF (1139 ms) (F(1,39)=22.1; p<.001; F(2,111)=26.1; p<.001).

No interaction was found in the field/sentence type analysis either (Fs < 1). There were main effects of field (F(1,39)=33.8; p<.001; F(2,111)=41.2; p<.001) and of sentence type (F(1,39)=55.7; p<.001; F(2,111)=142.9; p<.001). The first main effect reflected the slower processing of LVF stimuli (LVF: 1164 ms; RVF: 1066 ms), the second one represented the much faster times for the plausible sentences (implausible-related: 1199 ms; implausible-unrelated: 1178 ms; plausible-related: 968 ms).

Error Percentages The field/relatedness analysis did reveal significant interactions between field and relatedness in the error data (F(1,39)=4.3; p<.05; F(2,111)=3.9; p=.05). The semantic illusion effect turned out to be present in the LH, but totally absent in the RH: when target words were presented to the RVF, the amount of errors rose from 5 % (implausible-unrelated) to 9 % (implausible-related), as opposed to a null-result for targets presented to the LVF (both conditions 8 % errors). There were no main effects of field and relatedness (all p's > .20).

The field/sentence type analysis produced a marginally significant interaction of field and sentence type (F(1,39)=3.7; p=.06; F(2,111)=3.7; p=.06), essentially reflecting a trend toward a null effect of sentence type in the RH (stimuli presented to LVF: all

![Figure 3. Reaction times (upper panel) and Error percentages (lower panel) for Experiment 2. P-R=plausible-related; I-R=implausible-related; I-U=implausible-unrelated. LVF=left visual field; RVF=right visual field.](http://www.fineprint.com)
sentence types showed 8% errors) and a semantic illusion effect in the LH (stimuli presented to RVF: implausible-related 9%, both other sentences 5%). The main effect of field was significant in the analysis by items, but only marginally significant in the participants analysis (F(1,139)=3.5; p=.07; F(2,111)=4.0; p<.05) suggesting a trend of more errors when targets were presented in LVF (8%) as compared to RVF (6%). No main effect of sentence type was observed (all p's > .20).

**General Discussion**

The two behavioral experiments using the divided visual field paradigm have yielded two important findings. First, they provided convincing evidence for the existence of 'semantic illusions' during sentence processing. The experiments presented here replicated the original observation made by Hoeks et al. (2003), that sometimes semantic relationships between words in a sentence can be so strong, that they temporarily 'overrule' the semantic interpretation prescribed by the syntactic structure of that sentence. Experiment 1 using lexical decision showed that target words of implausible-related sentences were processed significantly faster and with fewer errors than the target words of implausible control-sentences. In addition, Experiment 2 showed that the semantic illusion effect tempts readers to classify the implausible-related sentences as plausible, causing significantly more errors there than in the implausible controls.

The second important finding is that although the results showed that semantic illusion does occur during sentence processing, there was only very little evidence for the presumed lateralization of the effect. Actually, the results of the two experiments taken together seem to implicate the LH as the most important site for the semantic illusion, against a null-effect in LVF/RH. Experiment 1 using reaction times again showed an RVF/LH focus for the presumed lateralization of the effect. Actually, the results of the two experiments taken together seem to implicate the LH as the most important site for the semantic illusion, against a null-effect in LVF/RH.

On the basis of these data it seems clear that we have to reject any model of sentence interpretation in which the LH is only focused on following the instructions of the solicited sentence context, and the RH is responsible for something else) and to look for a more dynamic model, in which the brain is expected to be very fruitful.

One possible way towards a solution might be to leave the 'static' models (e.g., LH is responsible for one thing and the RH is responsible for something else) and to look for a more dynamic model, in which the brain is expected to be very fruitful. Obviously, more research is needed to begin to answer these questions, despite the large number of studies that have already been done. But especially research in which divided visual field studies are combined with neuroimaging methods, such as ERPs and also fMRI (function magnetic resonance imaging) are expected to be very fruitful.

**Acknowledgements**

We are very grateful to Eran Zaidel and Berry Wijers for helpful comments on, and discussions of this work.

**References**


