Understanding words in sentence contexts: The time course of ambiguity resolution

Tamara Swaab,[a,*] Colin Brown,[b] and Peter Hagoort[b,c]

[a] *Department of Psychology, University of California, One Shields Avenue, Davis, CA 95616-8686, USA*
[b] *Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands*
[c] *F.C. Donders Centre for Cognitive Neuroimaging, Nijmegen, The Netherlands*

Abstract

Spoken language comprehension requires rapid integration of information from multiple linguistic sources. In the present study we addressed the temporal aspects of this integration process by focusing on the time course of the selection of the appropriate meaning of lexical ambiguities (''*bank*'') in sentence contexts. Successful selection of the contextually appropriate meaning of the ambiguous word is dependent upon the rapid binding of the contextual in-formation in the sentence to the appropriate meaning of the ambiguity. We used the N400 to identify the time course of this binding process. The N400 was measured to target words that followed three types of context sentences. In the concordant context, the sentence biased the meaning of the sentence-final ambiguous word so that it was related to the target. In the discordant context, the sentence context biased the meaning so that it was not related to the target. In the unrelated control condition, the sentences ended in an unambiguous noun that was unrelated to the target. Half of the concordant sentences biased the dominant meaning, and the other half biased the subordinate meaning of the sentence-final ambiguous words. The ISI between onset of the target word and offset of the sentence-final word of the context sentence was 100 ms in one version of the experiment, and 1250 ms in the second version. We found that (i) the lexically dominant meaning is always partly activated, independent of context, (ii) initially both dominant and subordinate meaning are (partly) activated, which suggests that contextual and lexical factors both contribute to sentence interpretation without context completely overriding lexical information, and (iii) strong lexical influences remain present for a relatively long period of time.
© 2003 Elsevier Science (USA). All rights reserved.

Keywords: Lexical ambiguities; Sentence context; Meaning frequency; ERPs; N400

1. Introduction

Lexically ambiguous words (e.g., *bank*) have the same form representation but two (or more) unrelated meanings. Many studies have been devoted to the resolution

---

of ambiguous words in sentence contexts (e.g., Conrad, 1974; Dopkins, Morris, & Rayner, 1992; Duffy, Morris, & Rayner, 1988; Glucksberg, Kreuz, & Rho, 1986; Hogaboam & Perfetti, 1975; Van Petten & Kutas, 1987; Lucas, 1987; Paul, Kellas, & Clark, 1992; Rayner & Frazier, 1989; Rayner, Pacht, & Duffy, 1994; Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982; Sereno, 1995; Simpson, 1981; Swaab, Brown, & Hagoort, 1998; Swinney, 1979; Tabossi & Zardon, 1993). The main focus of such studies has been to identify when context exerts its influence during normal language understanding: Immediately, at early levels of accessing representations in the semantic lexicon, or later during integration of syntactic and semantic information into a higher order message representations. Modular theories maintain the view that the initial activation of word meanings is not influenced by context, and that either multiple meanings of the ambiguous words are accessed (multiple or exhaustive access model: e.g., Conrad, 1974; Lucas, 1987; Onifer & Swinney, 1981; Swinney, 1979), or that the order of access is determined by the relative frequencies of the different meanings of the ambiguous words, with the most frequent meaning being retrieved first (ordered access model: e.g., Hogaboam & Perfetti, 1975).

In contrast, the "selective access" or "context dependent" models claim that lexical access can be influenced by context, and that as a consequence, only the contextually appropriate meaning of the ambiguous word is accessed when biased by the prior sentence context (e.g., Glucksberg et al., 1986; Simpson, 1981). These models are compatible with an interactive view of language processing (McClelland & Elman, 1986). In this view, the sensory input and the higher order representation of the linguistic context can interact at all stages of lexical processing, and hence, context can restrict access to the contextually appropriate meaning.

More recently, "constraint satisfaction" models have been proposed that maintain that certain principles that direct lexical processing in general also apply to the resolution of syntactic and semantic ambiguities. One of the basic principles of these models is that in order to understand a word in context (whether it is ambiguous or not), a single alternative has to be computed at each of three representational levels: the lexical level, the syntactic level, and the discourse level. MacDonald, Pearlmutter, and Seidenberg (1994) claim that this selection process can be described as a "constraint satisfaction problem" that is governed by lexical factors (like, for example, relative meaning frequency in the case of ambiguous words), and by contextual factors. Interactions between lexical and contextual factors will determine the product of the comprehension process, and contextual factors will play a role when the context provides information that helps to distinguish between alternatives. However, lexical factors like the frequency of the alternatives will limit the effects of contextual information. In the case of lexical (meaning) ambiguity resolution, this implies that context can influence the interpretation of equibiased ambiguities, but it "cannot overcome strong frequency biases to promote a subordinate interpretation over the (frequency) dominant alternative" (MacDonald et al., 1994, p. 693).

The results of a great number of behavioral studies that have used various measures (such as lexical decision, naming and eyetracking) in a variety of paradigms (including cross modal priming) indicate that ambiguity resolution in sentence contexts is influenced both by the degree of contextual constraint, and the relative frequency of the ambiguous words. Evidence in favor of multiple access modulated by frequency has been found when the sentence context is not too constraining (e.g., Rayner & Frazier, 1989). If the context contains words that are associatively related to one of the meanings of the ambiguous word, this meaning is activated or selected more rapidly (Seidenberg et al., 1982). If the context strongly biases toward the most frequent or dominant meaning of the lexical ambiguity by priming a salient feature of this meaning, then the less frequent or subordinate meaning might not be accessed

at all (Tabossi, 1988; Tabossi, Colombo, & Job, 1987). In contrast, the dominant meaning seems to always be accessed, even in contexts that strongly bias the subordinate meaning. Finally, some studies have shown that the time course of lexical selection may be different for the left than for the right hemisphere, with longer activation of both the contextually appropriate and the contextually inappropriate meaning in the right hemisphere (e.g., Burgess & Simpson, 1988; Coney & Evans, 2000; Swaab et al., 1998).

The present study used ERPs to further explore if and when lexical factors such as relative meaning frequency of ambiguous words influence real-time sentence comprehension, even when the sentence context clearly biases one of the meanings of an ambiguous word. Although behavioral studies strongly, but not exclusively, suggest this is so, in order to fully establish this mechanism it is essential that evidence from different methods and experimental approaches point to the same result. This is not now the case, as contradictory findings come from an ERP study that looked at the time course of ambiguity resolution in sentence contexts in young adults (Van Petten & Kutas, 1987). In this study the authors argued for the idea that a backward priming mechanism accounted for the activation of contextually inappropriate meanings of ambiguous words (Van Petten & Kutas, 1987). This means that the activation of the contextually inappropriate meaning was thought to occur as a function of the processing of the meaning of the target word, and not as a result of initial multiple activation. A backward priming account is at odds with the idea that lexical frequency can influence real-time sentence processing, and one of the goals of the present ERP study is to address this issue. Specifically, we investigated whether possible influences of lexical factors on real-time sentence comprehension were restricted to the more frequent or dominant meaning of the ambiguous words. In addition we were interested in identifying what the time course of potential lexical influences on sentence interpretation may be, and in particular, if such influences would be found after a relatively long interval of time (1250 ms).

ERPs were recorded to target words that followed three types of context sentences after an interval of silence (ISI). For the purposes of the present study, the N400 was the most relevant ERP component because it is especially sensitive to lexical integration processes (e.g., Brown & Hagoort, 1993; Holcomb, 1993; Rugg, Furda, & Lorist, 1988), and lexical integration is crucial to the process of contextual selection. In the present experiment, the sentences were presented as naturally produced connected speech, followed by the target words, also presented in the auditory modality. The ISI between onset of the target word and offset of the sentence-final word of the context sentence was 100 ms in the first version of the experiment, and 1250 ms in the second version of the experiment. In the concordant context, the sentences biased the meaning of the sentence-final ambiguous word that was related to the target. In the discordant context, the alternative meaning of the sentence-final ambiguous word was biased, which was unrelated to the target. In the unrelated control condition, the sentences ended in an unambiguous noun that was unrelated to the target. Half of the concordant sentences biased the dominant meaning, and the other half biased the subordinate meaning of the sentence-final ambiguous words.

Given the findings in the literature, we expected context effects at both the short and the long ISI for the contextually appropriate meaning, which is dominant in half of the sentences, and subordinate in the other half. This should result in a reduced N400 to targets in the concordant context relative to the unrelated and the discordant conditions at both ISIs. At the short ISI, we predicted that activation of the contextually inappropriate meaning is modulated by frequency, which would result in a stronger reduction of the N400 amplitude to targets in the discordant relative to

the unrelated condition for the dominant compared to the subordinate meaning condition. At the long ISI, the hypothesis was that the inappropriate meaning for both the subordinate and the dominant meanings of the lexical ambiguity should decay or be completely suppressed, and this would result in the absence of an N400 difference between the discordant and the unrelated targets.

2. Method

2.1. Participants

A group of 45 college-aged students of the University of Nijmegen were paid for their participation; 24 in the long ISI version of the experiment, and 21 in short ISI version. All participants were right handed according to an abridged version of the Oldfield Handedness Inventory (Oldfield, 1971).

2.2. Stimuli

Selection and construction of a subset of the materials has been described elsewhere (Swaab et al., 1998). For the purpose of the present study we will reiterate the relevant aspects of the materials here. Ambiguous nouns (114) with a dominant and a subordinate meaning were selected from 442 normed Dutch words with two or more unrelated meanings. For each ambiguous noun, a set of three context sentences was constructed, followed by a target word[1] (see Table 1—literal English translation of Dutch materials is provided).

Two of the three sentences ended with an ambiguous noun, and the third with an unambiguous noun. Sentences that ended in an ambiguous noun either biased the subordinate or the dominant meaning of this word. In the dominant condition, the target words were related to the dominant meaning of the sentence final ambiguous word, and in the subordinate condition to the subordinate meaning. As can be seen in the example, each set of three context sentences had the same constituent structure. Importantly, the content words in the sentences were never associatively related to the sentence-final word or the target presented after the sentence. Furthermore, *no* salient features of the meaning of the ambiguity were primed by the sentence context. Thus, for example, if the sentence ended in the sentence final ambiguous word "*bank*," words like "*rich*" and "*businessman*" that prime salient features of the "*money*" meaning of this ambiguous word were avoided.

2.2.1. Pretests
The set of 342 sentences (3 × 114) was submitted to a cloze test with 64 paid participants and an association test with 40 new subjects. None of the participants in the pre-tests took part in the ERP experiment. In the cloze test, subjects were presented with all the sentences in random order, each with the sentence final word omitted. They were asked to fill in the word that best fit the sentence context. The results of the cloze tests were used to establish that the maximum cloze probability did not exceed 45% (to avoid pre-selection of the word), and to match the mean cloze probability between conditions.

---

[1] Even though participants did not have to respond to the words that followed the sentences in the three context conditions, these words will be called targets in this study. This is done in order to be consistent with the nomenclature that is used in the literature.

Table 1
Examples of sentences in the six context conditions. The sentence-final words and the targets are in capitals

Context condition	Sentence	Target word
Target: related to dominant meaning		
Concordant	Mijn vriendin overlegt met haar ADVOCAAT (My friend meets with her LAWYER	RECHTER JUDGE)
Discordant	Mijn moeder lepelt van haar ADVOCAAT (My mother eats [of] her EGGNOG	RECHTER JUDGE)
Unrelated	De jongen ruziet over zijn SALARIS (The boy argues about his SALARY	RECHTER JUDGE)
Target: related to subordinate meaning		
Concordant	De chirurg verwijdert kundig de pijnlijke AMANDELEN (The surgeon removes skillfully the painful TONSILS	KEEL THROAT)
Discordant	De jongen pikt snel de grootste AMANDELEN (The boy takes quickly the biggest ALMONDS	KEEL THROAT)
Unrelated	De krant publiceert vandaag deze knappe SPOTPRENTEN (The newspaper publishes today these good CARTOONS	KEEL THROAT)

The association test followed the cloze procedure. It was performed to establish whether the sentence contexts were successful in biasing the intended readings of the sentence-final ambiguous nouns. In addition, for the unambiguous nouns in the unrelated condition the association test was used to establish that the target word was never given as a response. In this test the complete context sentences were presented to the participants, and they were required to read the sentence and to write down the first word that came to mind upon reading the sentence-final word. The two sentences ending in the same ambiguous word were assigned to two different versions of the association test. In this way none of the participants saw a sentence-final word more than once. The results of the association test were used to establish that the sentence contexts clearly biased the intended reading of the sentence final ambiguous word: at least 60% of the answers were related to the biased reading, and no more than 25% to the unbiased reading. After the pre-tests 100 sentence triplets were selected.

2.2.2. Selected materials

Targets words were chosen from the set of normed words that was described at the beginning of this materials section. For 50 of the 100 selected ambiguous nouns, strong associates of the dominant meaning were chosen as a target, for the remaining 50 strong associates of the subordinate meaning were chosen. These target words were combined with the sentence triplets to form the dominant and subordinate conditions. The dominant meanings of the ambiguous nouns in this set of items had an average frequency of occurrence of 85% (range 60–99%), the average frequency of occurrence for the subordinate meaning was 19% (range 2–40%). All sentences had a mean length of 8.6 words (range 6–13 words).

The results from the cloze tests revealed that the mean cloze probability for the sentence-final nouns was less than 7% (range 0–34%). None of the target words were mentioned more than once in these tests. The strength of the contextual bias for the selected sentences was computed from the results of the association tests. For the *dominant* items, on average 90% of the responses to the concordant context sentences was related to the intended dominant meaning, and 3% to the unintended subordinate

meaning. For the discordant context sentences in this condition, a mean of 83% was related to the intended subordinate meaning, and 8% to the unintended dominant meaning. For the *subordinate* items the contextual bias was as follows: in the concordant condition, on average 90% of the responses in the association test was related to the intended subordinate meaning, and 3% to the unintended dominant meaning. In the discordant context 85% of the responses was related to the intended dominant meaning, and 7% to the unintended subordinate meaning. These results indicate that the context sentences were successful in biasing the intended meaning of the ambiguous sentence-final nouns. Moreover, the bias was comparable in strength for the different conditions. Finally, the target word was never given as a response in the unrelated control sentences.

In addition to the 100 sentence triplets, 132 filler sentences were constructed. These sentences ended in an unambiguous word that was unrelated to the target word that followed it. In the combined set of experimental triplets and filler sentences, less than half of the sentence-final words was related to the target, which made the development of strategies to predict a relation between sentence-final word and target unlikely[2]. Finally, 19 practice sentences were included, half of which were followed by a related target, and half by an unrelated target.

All materials were spoken at normal rate and with normal intonation by an experienced female speaker in a sound-attenuating booth. The sentences and the targets were recorded on a digital audio tape and stored on a hard disk. A speech waveform editing system was used to add targets to their respective sentences. Identical targets were represented by the same physical token. The same materials were used in the two ISI versions of the experiment. In version one, the ISI was 100 ms, in version two it was 1250 ms. The inter-trial-interval was 6415 ms in both versions of the experiment.

Digital audio-tapes were constructed that contained the sentence materials. In addition a questionnaire was prepared with questions about the sentence content of 15 sentences.

2.3. Procedure

Two groups of participants were tested in sessions of about 4 h each, 24 in the long and 21 in the short ISI version of the experiment. In each ISI version of the experiment, experimental sentences were randomly intermixed with filler sentences, and these materials were divided over three experimental blocks. Every block began with four fillers. Targets and ambiguous words were never repeated within a block, and none of the sentences were repeated.

The presentation order of the blocks was counterbalanced, resulting in three presentation versions. Participants were assigned to one of three presentation versions such that in the short ISI version of the experiment, seven of the 21 participants first heard the stimuli in block 1, then in 2 and then in 3, seven were assigned to the second presentation version and they first heard block 2, then 3 and then 1, and seven were assigned to the third presentation version, and these participants first heard block 3, then 1, and then 2. In the long ISI version, 8 participants were tested in each presentation version.

Participants were tested individually in a dimly illuminated sound-attenuating booth, seated in a comfortable reclining chair. They were instructed to move as little as possible, and to keep their eyes fixated on a point at eye-level 1.5 m in front of

---

[2] Furthermore, given the fact that participants had no task that urged them to generate a fast response, it is unlikely that they developed a response-related strategy.

them. The participants were told that they would hear sentences followed by a word. They were asked to listen attentively to the materials, and they were told that the experimenter would sometimes stop the tape to ask them a question about a sentence they had just heard. No additional tasks were imposed.

All the stimuli were presented via a DAT-recorder (SONY 300ES). The participants listened to the stimuli via a closed-ear Sennheiser HMD-224 headphone. The trigger pulse on the right channel of the test tapes was inaudible to the participants.

In the 100 ms ISI version, the total run time (time during which stimuli were actually being presented) was 82 min including the practice list (3 min) and the three experimental blocks (each 24.5 min). In the 1250 ms ISI version, the total run time was about 91 min, including the practice list (3 min) and the three experimental blocks (each 27 min). Each session started with the practice list to familiarize the participants with the stimulus situation, and to train them to fixate their eyes during sentence presentation. Every 4–6 min, a question was asked about both the sentence contents of a control sentence, and about the target word that followed this sentence. This was done to verify that the participants were paying attention to the sentences and the target words. The answer was written down by the experimenter. If necessary, a short break was given after asking a question. All the participants were given a break of about 10 min after each experimental block.

2.4. EEG-recording

The electroencephalogram (EEG) was recorded from seven tin electrodes embedded in an elastic electrode cap, each referred to the left mastoid. Three electrodes were placed according to the International 10–20 system (Jasper, 1958) over midline sites at Fz, Cz and Pz locations. Four electrodes were placed laterally over symmetrical anterior and posterior positions. The symmetrical anterior electrodes (AL, AR) were located one half of the distance between F7 and T3, and F8 and T4 sites. The symmetrical posterior electrodes (PL, PR) were located 30% of the interaural distance lateral to a point 13% of the nasion-inion distance posterior to Cz. Vertical and horizontal eye movements were monitored by recording the electrooculogram (EOG) via sub- and supra-orbital electrodes, and left and right external canthus montages, respectively. The ground electrode was placed on the forehead, 10% from the nasion–inion distance above the nasion.

The EEG and EOG recordings were amplified by Nihon Kohden AB601G bio-electric amplifiers. EEG was recorded using a time constant of 8 s and hi-pass cut-off of 30 Hz. The electrode impedances were kept below 5 kΩ. The signal was digitized on-line with a sampling frequency of 200 Hz. The trigger pulse started sampling 150 ms before the presentation of a sentence. The total sampling epoch per trial was 8400 ms for the stimuli in the 1250 ms ISI version and 7400 in the 100 ms ISI version. Data were stored on a hard disk along with condition codes for off-line signal averaging and data analysis.

3. Results

Prior to off-line averaging, all single trial waveforms were screened for electrode drifting, amplifier blocking, muscle artifacts, eye movements, and blinks. This was done over an epoch of 2000 ms, starting 150 ms before the onset of the relevant words in the experimental sentences (i.e., the target words in the different context conditions). Trials containing artifacts were rejected. However, in some participants there were a substantial number of blinks contaminating the EEG data. In those

subjects, eye-blink artifacts were estimated and removed from the EEG prior to averaging using the procedure described by Gratton, Coles, and Donchin (1983). The overall rejection rate was 23% in the 100 ms ISI version of the experiment, and 25% in the 1250 ms ISI version of the experiment.

Repeated measures analysis of variance (ANOVA) were performed in the relevant latency ranges on the mean amplitudes of the N400 to the targets in the six context conditions. As has been observed previously (e.g., Van Petten & Kutas, 1987), the N400 effects showed a clear posterior distribution, and hence, the ANOVAs were done on the values from the centroparietal electrode sites (Cz, Pz, PL, and PR) where the N400 effects were of maximal amplitude. To compensate for inhomogeneous variances and covariances across treatment levels, the degrees of freedom for the F tests were adjusted using the procedure described by Greenhouse and Geisser (1959). The adjusted degrees of freedom and p-values will be presented.

Average waveforms to the targets in each of the six conditions were computed for each participant. First, overall repeated-measures ANOVAs were done on the mean amplitude of the N400 in the 300–600 ms epoch after target onset, relative to a 100 ms pre-target baseline. This epoch was selected to be consistent with previous literature on the N400 and after visual inspection of the waveforms. The ANOVAs had the following within-subject factors and levels: Frequency (related to the frequent (dominant) meaning, or the infrequent (subordinate) meaning of the ambiguous word), Congruency (concordant, discordant, or unrelated), and Electrode Site (Cz, Pz, PL, PR). In addition, a priori planned comparisons were conducted for the dominant and the subordinate targets separately, comparing the concordant relative to the unrelated condition, the discordant relative to the control condition, and the concordant relative to the discordant condition. These latter analyses had Congruency (concordant, discordant, or unrelated) and Electrode Site (Cz, Pz, PL, PR) as the within-subject factors.

3.1. Short ISI

The averaged waveforms in the three context conditions for the dominant targets are shown in Fig. 1, and for the subordinate targets in Fig. 2.

The waveforms in Figs. 1 and 2 have the following characteristics. There are small, but clearly identifiable auditory-evoked N1 and P2 components. The amplitudes of these components are probably slightly attenuated because the ISI between sentence-final word and target was only 100 ms, and these responses are known to be sensitive to ISI, presumably due to neural refractoriness (e.g., Knight, Hillyard, Woods, & Neville, 1980). Overall, the P2 is larger in amplitude in the concordant and discordant conditions than in the unrelated condition. However, this difference might actually reflect differences in the amplitude of the leading edge of the subsequent N400 response (e.g., Hagoort, Brown, & Swaab, 1996). The N1 and P2 components are followed by the N400, which has a clear centroparietal distribution. The amplitude of the N400 to targets in the concordant condition is reduced when compared to the unrelated and the discordant conditions for the dominant and the subordinate targets. The amplitude of the N400 to targets in the discordant condition is also reduced relative to the unrelated condition. This latter difference is larger for the dominant targets. The N400 is followed by a positive deflection over the posterior leads, most evident for the targets in the concordant condition. The waveform for the unrelated targets remains more negative than the waveforms for the discordant and the concordant targets, while the waveforms for the targets in the concordant and the discordant condition start to overlap at the end of the measurement window.
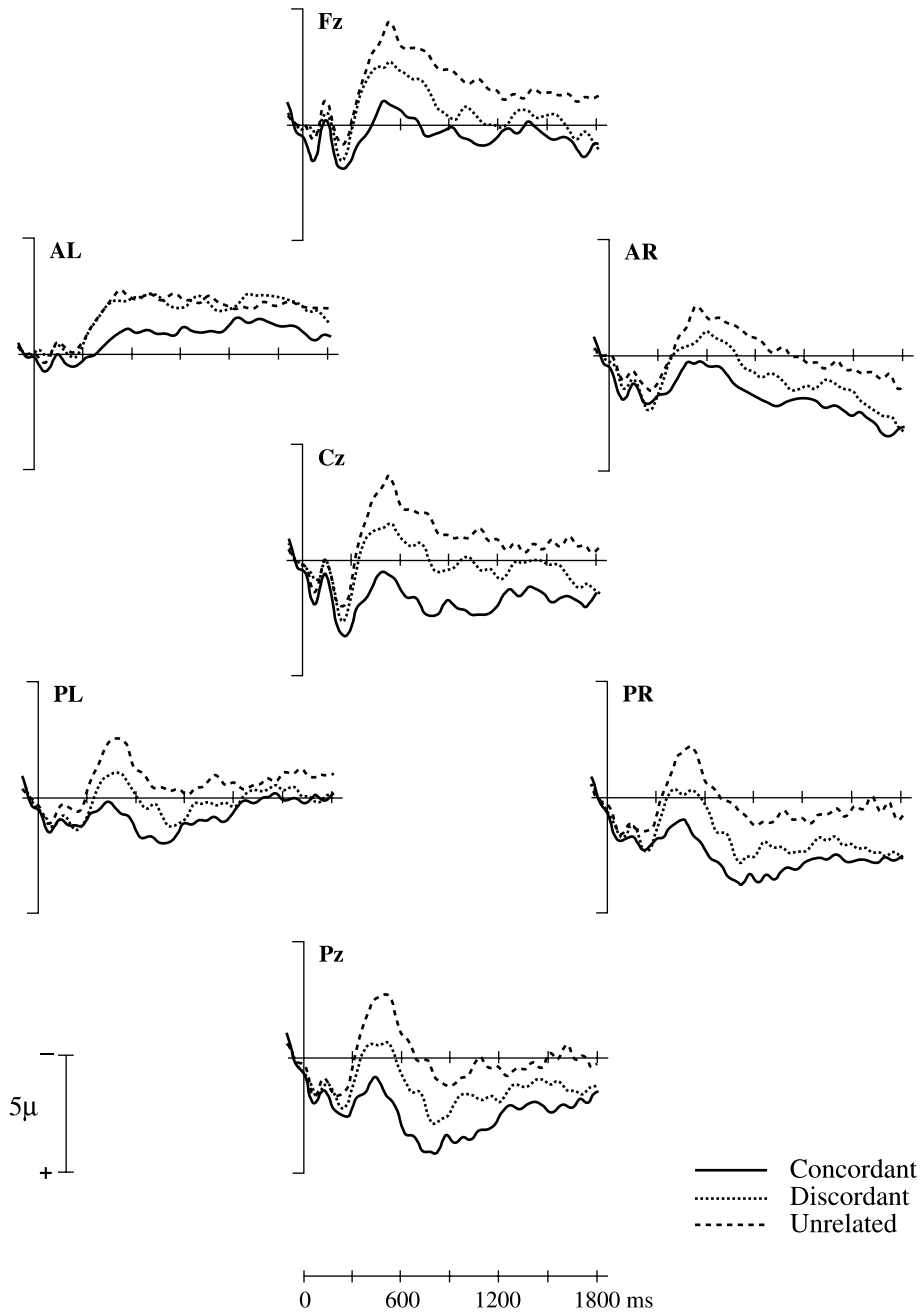
ISI = 100 ms, Dominant Targets

Fig. 1. Grand average ERPs in the short (100 ms) ISI version for the dominant targets in the concordant (solid line), discordant (dotted line), and unrelated (dashed line) context conditions.

An overall ANOVA was done on the N400-amplitude for all six conditions in the 300–600 ms epoch. This ANOVA had Order of Presentation (Block 1,2,3; Block 2,3,1; Block 3,1,2) as the additional factor, with participants nested under Order of

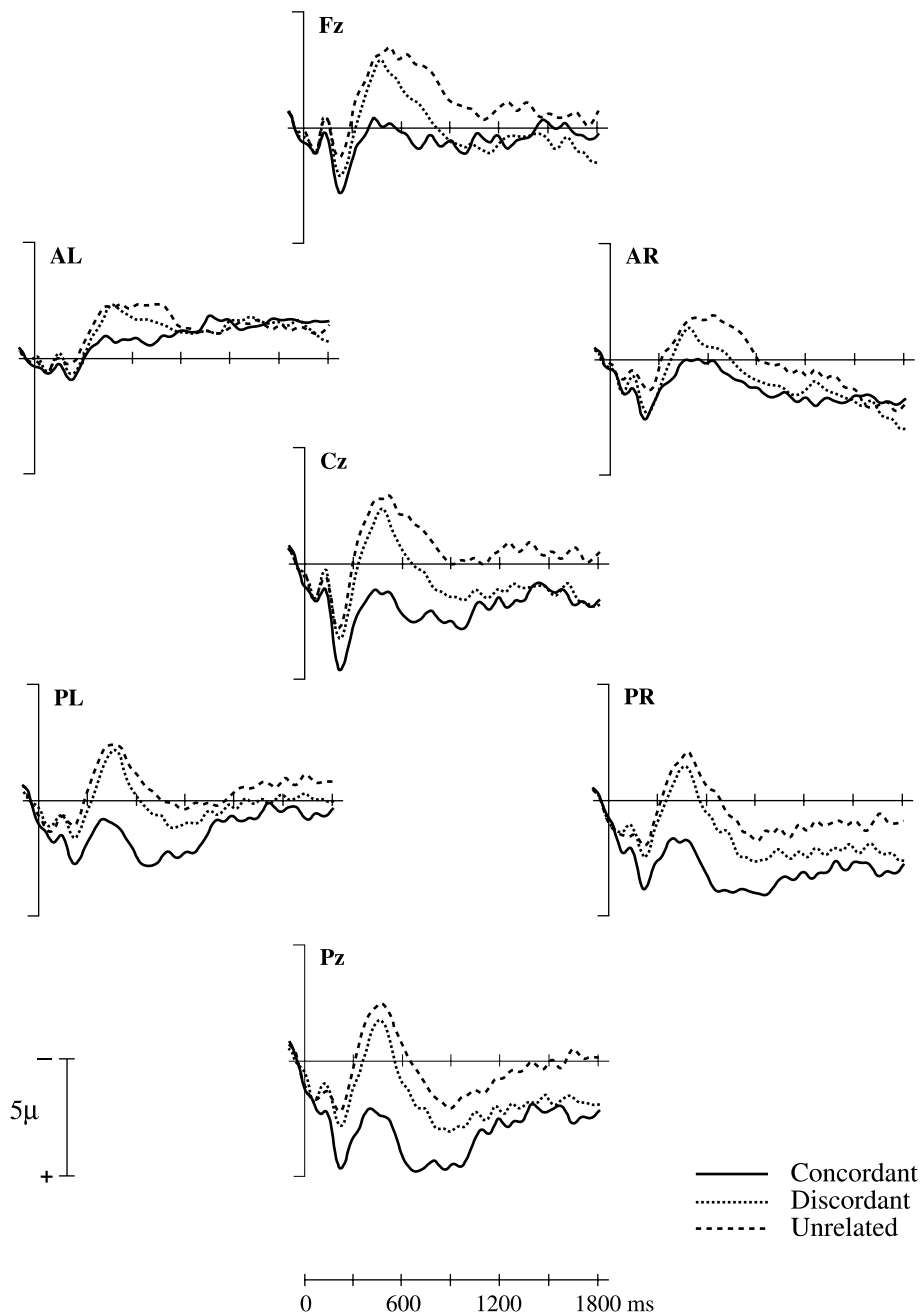ISI = 100 ms, Subordinate Targets

Fig. 2. Grand average ERPs in the short (100 ms) ISI version for the subordinate targets in the concordant (solid line), discordant (dotted line), and unrelated (dashed line) context conditions.

Presentation. There was no main effect of Order of Presentation ($F < 1$), and no interaction of Order of Presentation with Congruency ($F < 1$) or Frequency ($F < 1$), indicating that the Order of Presentation of the blocks had not influenced the pattern

of results. The ANOVA yielded no main effect of Frequency ($F < 1$). The amplitude of the N400 was comparable to targets related to the dominant and targets related to the subordinate meaning of the ambiguous word. The difference between conditions was substantiated by a significant main effect of Congruency ($F(1, 18) = 68.47$, $p < .01$). The effect of Congruency did not interact with Frequency ($F(1, 18) = 4.01$, $p > .05$).

In a series of a priori planned comparisons, the differences between the three context conditions were tested for the dominant and the subordinate targets separately. As expected, for the *dominant targets* there was a significant reduction of the amplitude of the N400 for the targets in the concordant relative to the unrelated condition ($F(1, 18) = 78.11$, $p < .0001$; $-3.03 \,\mu V$), and for the concordant relative to the discordant condition ($F(1, 18) = 26.37$, $p < .0002$; $-1.60 \,\mu V$). There was also a highly significant reduction of the N400 amplitude to the targets in the discordant relative to the unrelated condition ($F(1, 18) = 39.13$, $p < .0001$; $-1.43 \,\mu V$).

For the *subordinate targets* there was a significant reduction of the N400 amplitude to the targets in the concordant relative to the unrelated condition ($F(1, 18) = 50.39$, $p < .0001$; $-3.65 \,\mu V$), and also for the targets in the concordant relative to the discordant condition ($F(1, 18) = 48.11$, $p < .0001$; $-2.78 \,\mu V$). The relatively small N400 difference between the discordant and the unrelated condition was also highly significant ($F(1, 18) = 9.75$, $p < .006$; $-0.87 \,\mu V$).

In sum, both dominant and subordinate targets showed a significant reduction of the N400-amplitude in the concordant condition relative to the other two context types. Furthermore, the N400-amplitude was significantly reduced to both target types in the discordant relative to the unrelated condition. The reduction in the discordant context was however much smaller than the reduction in the concordant context. In addition, the difference between the discordant and the unrelated condition was larger for the dominant targets than for the subordinate targets by $0.56 \,\mu V$.

3.2. Long ISI

The analyses of variance for this version of the experiment were performed according to the procedures that were described in Section 3.1 for the short ISI.

The averaged waveforms in the six context conditions are shown in Fig. 3 for the dominant targets, and in Fig. 4 for the subordinate targets.

As can be seen in Figs. 3 and 4, with an ISI of 1250 ms the waveforms contain clear N1 and P2 components that are largest over the anterior and central sites. The N1 and P2 components are followed by the N400, that is larger over centroparietal sites. The amplitude of the N400 is reduced in the concordant relative to the unrelated and the discordant conditions. For the dominant targets, over the centroparietal sites the N400 in the discordant condition is reduced relative to the unrelated condition. The difference between the unrelated and discordant conditions is smaller or absent for the subordinate targets. The N400 is followed by a positive deflection over the posterior leads in the concordant condition, and for the dominant targets this positive deflection is also seen in the discordant condition. For both the dominant and the subordinate targets, the waveforms for the concordant and the discordant conditions start to overlap in the later part of the measurement window, while the waveform in the unrelated condition remains more negative.

An overall ANOVA on the mean amplitude of the N400 in the 300–600 ms epoch revealed no main effect of Order of Presentation ($F < 1$), nor did this factor interact with Congruency ($F < 1$) or Frequency ($F < 1$). This pattern indicates once again that Order of Presentation did not influence the results. Just as in the short ISI
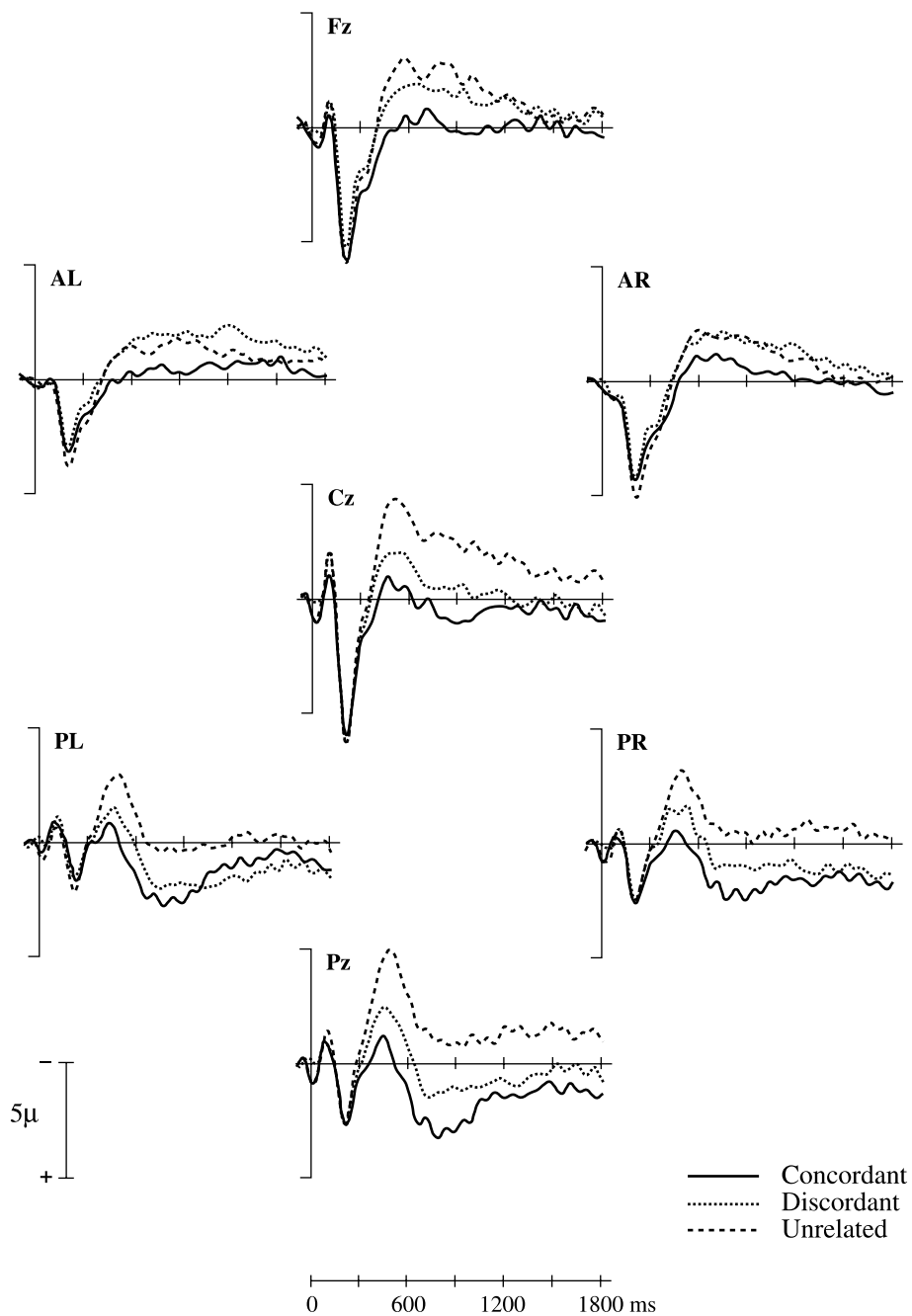
ISI = 1250 ms, Dominant Targets

Fig. 3. Grand average ERPs in the long (1250 ms) ISI version for the dominant targets in the concordant (solid line), discordant (dotted line), and unrelated (dashed line) context conditions.

version, there was a significant main effect of Congruency ($F(1, 21) = 40.21$, $p < .01$), and no significant Frequency by Congruency interaction ($F(1, 21) = 1.91$, $p > .05$). But, in contrast to the short ISI version, there was a significant effect of
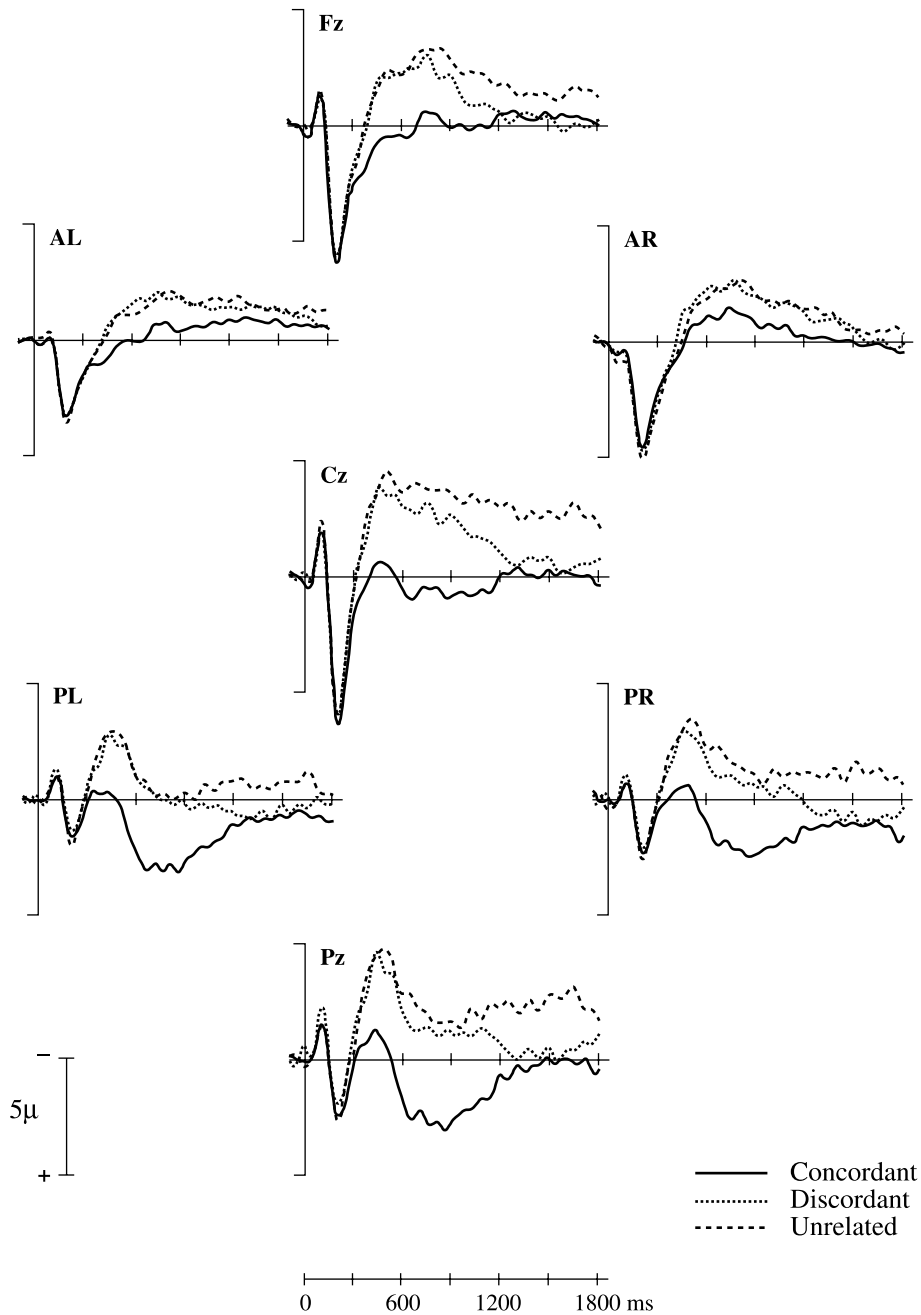
ISI = 1250 ms, Subordinate Targets

Fig. 4. Grand average ERPs in the long (1250 ms) ISI version for the subordinate targets in the concordant (solid line), discordant (dotted line), and unrelated (dashed line) context conditions.

Frequency ($F(1, 21) = 6.29$, $p < .03$). Inspection of the waveforms in Figs. 3 and 4 reveals that this was due to the fact that the amplitude of the N400 was somewhat larger to the subordinate than to the dominant targets in all context conditions.

The planned comparisons for the *dominant targets* revealed a significant reduction of the N400 amplitude in the concordant relative to the unrelated condition ($F(1, 21) = 16, 72, p < .0007$; $-2.44\,\mu V$), in the concordant relative to the discordant condition ($F(1, 21) = 8.43, p < .009$; $-1.08\,\mu V$), and in the discordant relative to the unrelated condition ($F(1, 21) = 6.16, p < .03$; $-1.36\,\mu V$).

For the *subordinate* targets, there was also a significant N400 amplitude reduction for the targets in the concordant condition relative to the unrelated condition ($F(1, 21) = 29.28, p < .0001$; $-2.74\,\mu V$), and for the targets in the concordant relative to the discordant condition ($F(1, 21) = 61.50, p < .0001$; $-2.44\,\mu V$). No significant difference was obtained between the unrelated and discordant conditions ($F < 1$; $-0.30\,\mu V$).

In summary, in the long ISI version the N400-amplitude was significantly reduced to both the dominant and the subordinate targets in the concordant condition relative to the two other context types. In addition, there was a reduction of the amplitude of the N400 to the dominant targets in the discordant relative to the unrelated condition, but this was not the case for the subordinate targets. In other words, there was evidence of activation of the dominant meaning when it was contextually inappropriate. In contrast, the subordinate meaning had decayed or was suppressed when it was contextually inappropriate[3].

4. Discussion

Because ambiguous words have the same word form, but more than one distinct meaning (e.g., "*bank*"), they confront the listener with the problem of selecting only the meaning that matches with the preceding context. In the present study, we used the N400 to look at the time course of this contextual selection process. Specifically, the present experiment explored whether and when lexical factors (i.e., meaning frequency) interact with contextual constraints provided by sentence context during spoken sentence comprehension. The results indicate that lexical factors do influence sentence comprehension, even given a relatively long period of time after sentence completion (i.e., 1250 ms). We will discuss these results in the light of previous findings, and then present our conclusions.

Ambiguity resolution in sentence contexts has previously been investigated using ERPs by Van Petten and Kutas (1987). They asked participants to read sentences that biased the subordinate meaning of the sentence-final ambiguous word. The sentences were followed by a visually presented target after a 216 or 700 ms SOA. As expected, they found a reduction of the N400 to the contextually appropriate targets (related to the subordinate meaning) relative to the unrelated targets in both SOA conditions. The pattern of results for the contextually inappropriate targets (related to the dominant meaning) was, however, less straightforward. In their long SOA condition, the N400 to the contextually inappropriate targets was *not* different from the N400 to the unrelated targets. Initially, there was also no N400 difference between the unrelated and the contextually inappropriate targets in the short SOA condition, but after 500 ms the waveforms started to diverge, and the waveform to the contextually inappropriate targets became more positive than the waveform to

---

[3] To check whether repetition of targets and ambiguous words affected the results of this study, analyses were done on the data from the block that was first presented to each participant. Targets and ambiguous words were never repeated within a block. Nevertheless, all the experimental items were included in this analysis, because each of the three blocks occurred first in one of the three presentation versions. Absence of effects of repetition can be inferred if the pattern of results for this first block analysis closely matches the overall analysis, and this was indeed the case.

the unrelated targets. Van Petten and Kutas argued that this pattern of results could *not* be explained in terms of multiple access, because this would require "simultaneous and parallel activation of both senses of the ambiguous word" (Van Petten & Kutas, 1987, p. 200), and they had found a 200 ms delay in the effect for the contextually inappropriate targets relative to the effect for the contextually appropriate targets. Instead they proposed that the late effect for the contextually inappropriate targets resulted from temporal backward priming where the processing of the target influenced the processing of the prior prime (Kiger & Glass, 1983).

In the present study however, there was no evidence for the idea that any of the effects could be attributed to temporal backward priming: The difference in the N400 amplitude to targets in the discordant and the unrelated contexts was apparent over the whole 300–600 ms N400 time window[4]. Several methodological differences could have contributed to the discrepancy between the present study and that of Van Petten and Kutas (1987), including different modalities of presentation (visual in Van Petten and Kutas, auditory in the present study), and different types of contextual constraint: Van Petten and Kutas included semantically related words in their sentence contexts to bias the subordinate meaning of the ambiguous words, whereas in the present study such semantic relations were avoided. But importantly, regardless of remaining uncertainties as to the reasons for differences between the two studies, the results of the present study cannot easily be explained in terms of backward priming.

In general, the pattern of results of the present study is consistent with the idea that both the dominant and the subordinate meanings of ambiguous words were activated, regardless of the bias of the sentence context. For both the dominant and the subordinate conditions the N400 was reduced to the targets in the concordant relative to the unrelated condition, which indicates that the appropriate meaning of the ambiguous word had been integrated[5]. At the short ISI, there was also evidence that the contextually inappropriate meanings had been activated in both conditions, because there was a reduction of the amplitude of the N400 to the targets in the discordant relative to the unrelated condition. This effect was larger for the dominant than for the subordinate meaning of the ambiguous words. Nevertheless, even subordinate meanings of ambiguous words were activated when they were contextually inappropriate. This finding was not expected, given that normal elderly control subjects do not show activation of contextually inappropriate subordinate meanings of ambiguous words (Swaab et al., 1998). Recently, a similar pattern of results was reported for noun–verb ambiguities; i.e., activation of the irrelevant subordinate meaning of ambiguous words for young but not for older adults (Winkler & Swaab, 2001). These findings are consistent with the idea that older adults make better use of contextual constraints than young adults, possibly because they have had more language experience.

---

[4] In order to further test the onset of the effects in the different conditions we performed multiple ANOVA's over 25 ms epochs in the 300–600 ms N400 time window. These analyses also failed to show any evidence to support the idea that the discordant and unrelated conditions diverge from each other later than from the concordant condition (all onset at 300 ms). This supports the overall analyses over the 300–600 ms time-window, and argues against an interpretation in terms of backward priming.

[5] In a previous study, Hagoort and Brown (1994) have shown that the ERP waveform to ambiguous words is more negative than to unambiguous words, which raises a concern of differential overlap of components to the target words in the present study. However, in the present study the ERPs to the unrelated target words that follow unambiguous sentence final words were in most cases more negative in amplitude than ERPs to target words in the concordant and discordant conditions that follow ambiguous sentence final words, which is therefore in the opposite direction of what would be predicted from overlap of effects to sentence final words.

In the long ISI version, no evidence of activation of the contextually inappropriate subordinate meaning of the ambiguous word was found. However, even after this relatively long period of time (1250 ms), the dominant meaning of the ambiguous words was still partly active. Apparently, the inappropriate dominant meaning of ambiguous words remains available for a much longer period of time then was previously observed in behavioral studies. The absence of a task in the present study may have revealed properties of the influence of lexical frequency on real time sentence processing that are obscured in behavioral studies that require tasks because they introduce unwanted processes (e.g., subject strategies). However, this pattern of activation of the contextually irrelevant meanings does not indicate that context did not exert its influence at all, because in that case there would have been no difference between the N400 effects in the unrelated-concordant, and the unrelated-discordant conditions. This was clearly not the case: The differences between the unrelated and the discordant condition in the short ISI and for the dominant condition in the long ISI were small compared to the difference between the unrelated and the concordant conditions, which indicates that the sentences had been effective in engendering a bias. It is also important to note that there was a significant difference between the concordant and the discordant conditions, indicating that the activation of the contextually irrelevant meanings was clearly reduced.

One final aspect of the waveforms for both the subordinate as well as for the dominant targets should not remain unmentioned. Whereas the N400 to the targets in the concordant and the discordant conditions start to overlap in the late part of the measurement window in both ISI versions, the N400 to the targets in the unrelated condition remained more negative. It might be the case that the late difference between the discordant and the unrelated condition indicates the participants' late realization that the ambiguous word had an alternative meaning.

In conclusion, the results of the present study argue against a pure selective access model, but are consistent with both modular and constraint-based accounts of lexical ambiguity resolution, and support the idea that dominance plays a role in ambiguity resolution. Specifically, we found that the lexically dominant meaning is always partly activated, independent of context. Furthermore, we found clear evidence that, at least initially, both dominant and subordinate meaning are (partly) activated. This suggests that contextual and lexical factors both contribute to sentence interpretation without context completely overriding lexical information. And finally, in light of the fact that ERP effects show very rapid context effects (see also Van Berkum, Brown, & Hagoort, 1999), lexical influences, when they are strong, remain present for a relatively long time.

References

Brown, C. M., & Hagoort, P. (1993). The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience, 5*, 34–44.

Burgess, C., & Simpson, G. (1988). Cerebral hemispheric mechanism in the retrieval of ambiguous word meanings. *Brain and Language, 33*, 86–103.

Coney, J., & Evans, K. D. (2000). Hemispheric asymmetries in the resolution of lexical ambiguity. *Neuropsychologia, 38*, 272–282.

Conrad, C. (1974). Context effects in sentence comprehension: A study of the participative lexicon. *Memory & Cognition, 2*, 130–138.

Dopkins, S., Morris, R. K., & Rayner, K. (1992). Lexical ambiguity and eye fixations in reading: A test of competing models of lexical ambiguity resolution. *Journal of Memory and Language, 31*, 461–476.

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language, 27*, 429–446.

Glucksberg, S., Kreuz, R. J., & Rho, S. H. (1986). Context can constrain lexical access: Implications for models of language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*, 323–335.

Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology, 55*, 468–484.

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24*, 95–112.

Hagoort, P., & Brown, C. M. (1994). Brain responses to lexical ambiguity resolution and parsing. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Hagoort, P., Brown, C., & Swaab, T. Y. (1996). Lexical-semantic event-related potential effects in patients with left hemisphere lesions and aphasia, and patients with right hemisphere lesions without aphasia. *Brain, 119*, 627–649.

Hogaboam, T. W., & Perfetti, C. A. (1975). Lexical ambiguity and sentence comprehension. *Journal of Verbal Learning and Verbal Behavior, 14*, 265–274.

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

Jasper, H. H. (1958). Report to the committee on methods and clinical examination in electroencephalography. Appendix: The ten-twenty system of the International Federation. *Electroencephalography and Clinical Neuropsychology, 10*, 371–375.

Kiger, J. I., & Glass, A. L. (1983). The facilitation of lexical decisions by a prime occurring after the target. *Memory & Cognition, 11*, 356–365.

Knight, R. T., Hillyard, S. A., Woods, D. L., & Neville, H. (1980). The effects of frontal and temporal parietal lesions on the auditory evoked potentials in man. *Electroencephalography and Clinical Neurophysiology, 50*, 112–124.

Lucas, M. (1987). Frequency effects on the processing of ambiguous words in sentence contexts. *Language and Speech, 30*, 25–46.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review, 101*, 676–703.

McClelland, J. L., & Elman, J. L. (1986). Interactive processes in speech perception: The TRACE model. In J.L. McClelland, D.E. Rumelhart, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition*. Vol. 2. Cambridge, MA: MIT Press.

Oldfield, R. C. (1971). The assessment of the analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*, 97–113.

Onifer, W., & Swinney, D. A. (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. *Memory & Cognition, 9*, 225–236.

Paul, S. T., Kellas, M. M., & Clark, M. B. (1992). Influence of contextual features on the activation of ambiguous word meanings. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 703–717.

Rayner, K., & Frazier, L. (1989). Selection mechanisms in reading lexically ambiguous words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 779–790.

Rayner, K., Pacht, J. M., & Duffy, S. A. (1994). Effects of prior encounter and global discourse bias on the processing of lexically ambiguous words: Evidence from eye fixations. *Journal of Memory and Language, 33*, 527–544.

Rugg, M. D., Furda, J., & Lorist, M. (1988). The effects of task on the modulation of event-related potentials by word repetition. *Psychophysiology, 25*, 55–63.

Seidenberg, M. S., Tanenhaus, M. K., Leiman, J. M., & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge based processing. *Cognitive Psychology, 14*, 489–537.

Sereno, S. C. (1995). Resolution of lexical ambiguity: Evidence from an eye movement priming paradigm. *Journal of Experimental Psychology: Learning, Memory, and Behavior, 21*, 582–595.

Simpson, G. (1981). Meaning dominance and semantic context in the processing of lexical ambiguity. *Journal of Verbal Learning and Verbal Behavior, 20*, 120–136.

Swaab, T. Y., Brown, C., & Hagoort, P. (1998). Understanding ambiguous words in sentence contexts: Electrophysiological evidence for delayed contextual selection in Broca's aphasia. *Neuropsychologia, 36*, 737–761.

Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior, 18*, 645–659.

Tabossi, P. (1988). Accessing lexical ambiguity in different types of sentential contexts. *Journal of Memory and Language, 27*, 324–340.

Tabossi, P., Colombo, L., & Job, R. (1987). Accessing lexical ambiguity: Effects of context and dominance. *Journal of Memory and Language, 32*, 359–372.

Tabossi, P., & Zardon, F. (1993). Processing ambiguous words in context. *Journal of Memory and Language, 32*, 359–372.

Van Berkum, J. J. A., Brown, C. M, & Hagoort, P. (1999). Early referential context effects in sentence processing: Evidence from event-related potentials. *Journal of Memory and Language, 41*, 147–182.

Van Petten, C., & Kutas, M. (1987). Ambiguous words in context: An event-related potential analysis of the time course of meaning activation. *Journal of Memory and Language, 26*, 188–208.

Winkler, T., Swaab, & T. Y. (2001). Category ambiguity resolution in aging: An ERP study. *Journal of Cognitive Neuroscience*, (Suppl.), 165.