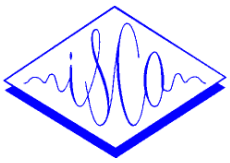

ISCA ITRW on
Spoken Word Access Processes (SWAP)
Nijmegen, The Netherlands
May 29-31, 2000

# ELECTROPHYSIOLOGICAL EVIDENCE FOR EARLY CONTEXTUAL INFLUENCES DURING SPOKEN-WORD RECOGNITION: THE N200

**D. van den Brink, C. M. Brown** and **P. Hagoort**
Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

## ABSTRACT

An event-related brain potential experiment was carried out to investigate the time-course of contextual influences on spoken-word recognition. Subjects were presented with spoken sentences that ended with a word that was either (a) congruent, (b) semantically anomalous beginning with the same initial phonemes as the congruent completion, or (c) semantically anomalous beginning with phonemes that differed from the congruent completion. In addition to finding an N400 component in the two semantically anomalous conditions, reflecting lexical-semantic integration processes, we obtained an early negative component in all conditions. This component, the N200, was largest in the semantically anomalous condition where word onset differed from that of the congruent completions. It was concluded that the N200 could be an indicator of the lexical selection process, where word-form information resulting from an initial phonological analysis, and content information derived from the context interact.

## 1. INTRODUCTION

In everyday speech, people hear words in the context of other words, usually in the form of sentences. In the literature on language comprehension, there is evidence to suggest that contextual influences play a role in the on-line recognition of spoken words. In the recognition of words spoken in isolation, a number of spoken-word recognition models converge on the idea that multiple lexical candidates that share word onset are accessed on the basis of an analysis of the initial phoneme(s) of a word ([1],[2],[3],[4]). As pronunciation of a word progresses over time, lexical candidates are dropped as soon as they no longer correspond to the incoming acoustic signal. Selection of the proper candidate is said to take place when only one candidate is left that still matches the acoustic signal.

Word recognition in sentences additionally requires that the selected word is integrated into a higher-order meaning representation of the sentence context. The impact of sentential-contextual information on the recognition process is a matter of debate, with several models assuming a high degree of interactivity between contextual and acoustic information, e.g. the TRACE model by McClelland and Elman [3], and other models assigning priority to the acoustic analysis, e.g. the Shortlist model by Norris [4] and the early and revised Cohort model by Marslen-Wilson ([1],[2]). A key issue here concerns the relative moment in time at which context starts to exert an effect on word recognition. The present study was designed to investigate the time-course of contextual effects on spoken-word recognition with event-related potentials (ERPs).

ERPs reflect the sum of simultaneous post-synaptic activity of a large number of neurons, recorded at the scalp as small voltage fluctuations in the electroencephalogram. A central finding in the ERP literature on language is a negative-going component that typically peaks at 400 ms after stimulus onset, the N400 ([5]). The N400 component is related to semantic processing of the eliciting word and is observed in both the visual and the auditory modality. It is believed that in sentence contexts, the N400 amplitude indexes the relative ease of semantic integration ([6]): Words that are incongruent or less fitting given the preceding sentence frame typically elicit a much larger N400 than words that fit well within the context. This N400 effect has a posterior scalp distribution and an onset around 200-250 ms after word onset (for a review, see [7]). However, semantic integration of the perceived word into the sentence context is one of the last subprocesses in spoken-word recognition. If the N400 solely reflects the semantic integration process, is there any other evidence in the electrophysiological signal that context exerts an influence prior to the semantic integration process? Or does the N400 component not only reflect integration processes, but also semantic processing at a level where lexical and contextual information interact?

The latest moment at which context starts to exert its influence was assumed to be revealed as the onset of a congruity effect in the waveforms, as indicated by the divergence between a congruent and an incongruent condition. We used semantically constraining Dutch sentences with sentence-final words that differed across three conditions. In the Fully Congruent condition sentences ended with the highest cloze probability word for that sentence: "De schilder kleurde de details in met een klein *penseel*" ("The painter colored the details with a small brush"). The other two conditions both ended anomalously, but the point at which the sentence-final words became incongruent differed between these two conditions. The completions of the sentences in the Initially Congruent condition (IC) began with the same phonemes as the highest cloze probability words in the Fully Congruent completions (FC): "De schilder kleurde de details in met een klein *pensioen*" ("The painter colored the details with a small pension"). In contrast, the Fully Incongruent sentence-final words (FI) had initial phonemes that differed from the highest cloze

probability word: "De schilder kleurde de details in met een klein *doolhof*" ("The painter colored the details with a small labyrinth"). We hypothesized that the onset of the congruity effect between the Fully Congruent and Fully Incongruent condition should precede the onset of the congruity effect between the Fully Congruent and Initially Congruent condition. We were also curious to see how this congruity effect would manifest itself; as the onset of a monophasic N400, indicating that the N400 also reflects contextual influences on word recognition (cf. [8]), or as a separate component preceding the N400, indicating a separate early process ([9], [10]).

## 2. METHODS

### 2.1. Subjects

The experiment was conducted with 21 native speakers of Dutch (18 female, mean age 22, range 19-26 years) from the subject pool of the Max Planck Institute for Psycholinguistics.

### 2.2. Materials

The experimental items consisted of a set of 261 triplets of semantically constraining sentences across the three experimental conditions mentioned above. The sentence-final words in IC had the same lexical stress and shared mean length of overlap of 2.6 phonemes, (range 2 to 4, always including a full vowel) with the FC completions. The sentence-final words had a mean duration of 525 ms, which did not differ across conditions. Because the sentences within each triplet were identical up to the final word, they had the same length (mean length of 10.8 words; range 5-15) and were equally constraining; the critical words in FC had a mean cloze probability of .84 (range: .50-1.0). All critical words had a frequency of at least 30 counts per 42 million ([11]).

### 2.3. Procedure

Subjects were tested individually in a dimly illuminated sound-attenuating booth. Their task was to attentively listen to the sentences and to try to understand them.

### 2.4. EEG-recording

The EEG was recorded from 29 silver-chloride electrodes mounted in an elastic cap, each referred to the left mastoid. The EEG and EOG recordings were amplified with a SynAmp Model 5083 EEG amplifier using a hi-cut of 30 Hz (notch filter 60 Hz) and a time constant of 8 s (0.02 Hz). Electrode impedances were kept below 3 kOhm. The EEG and EOG signals were digitized on-line with a sampling frequency of 200 Hz.

## 3. RESULTS

Figure 1 displays the grand average waveforms by electrode site time-locked to the onset of the sentence-final word. There are several things to note. First, the waveform of FI starts to diverge from the FC waveform earlier than the IC waveform. Second, a negativity at approximately 200 ms is visible in the waveforms. This negativity is apparent in all three conditions, but is largest in FI and is dominant over the frontal sites (see also Fz in Figure 2). Third, the semantically anomalous sentence-final words in IC and FI elicit a broad negativity peaking at approximately 400 ms, which is more negative than the ERP elicited in FC. In turn, this broad negativity is larger in FI than in IC. Its latency characteristics and morphology are similar to previously reported N400 effects.

Statistical analyses of the congruity effects consisted of a number of repeated measures analyses of variance (ANOVAs) with mean amplitude values computed for each subject and each electrode in two latency windows: (a) 150-250 ms after critical-word onset for the early negativity (N200), and (b) 300-500 ms after final-word onset for the N400. For each latency window, the results were analyzed by means of a priori pairwise comparisons between the congruity conditions using ANOVAs that crossed two levels of the congruity factor with the 29-level electrode factor. In addition to the pairwise comparisons, scalp distributions of the congruity effects were explored in separate ANOVAs.

*The Early Negativity/N200 latency window: 150-250 ms.* The a priori pairwise comparisons revealed that the FI completions elicited a larger N200 than the FC completions ($F[1, 20] = 5.70$, $MSe = 94.68$, $p < .05$, effect size = 0.56 μV), and that the ERPs in FC and IC did not significantly differ in amplitude ($F < 1$, effect size = 0.07 μV). In these pairwise analyses, none of the interactions of congruity with electrodes reached significance. Topographical analyses of FC versus FI revealed no significant interactions of congruity with site.

*The N400 latency window: 300-500 ms.* Statistical analysis revealed that both the FI and IC completions elicited a larger N400 than the FC completions ($F[1, 20] = 51.27$, $MSe = 3123.11$, $p < .001$ and $F[1, 20] = 23.87$, $MSe = 906.72$, $p < .001$, the effect sizes were 3.20 μV and 1.72 μV, respectively). Interactions with electrodes were also obtained in the pairwise comparisons (both $p < .001$). In addition, FI and IC differed significantly from each other in the N400 latency window ($F[1, 20] = 23.76$, $MSe = 664.24$, $p < .001$, effect size = 1.48 μ). An interaction with electrodes was also found ($p < .005$). Topographical analyses showed that the congruity effects in relation to the baseline were significantly larger over posterior than anterior regions of the scalp (FC vs. FI, $p < .001$; FC vs. IC, $p < .05$). No differences between hemispheres were found.

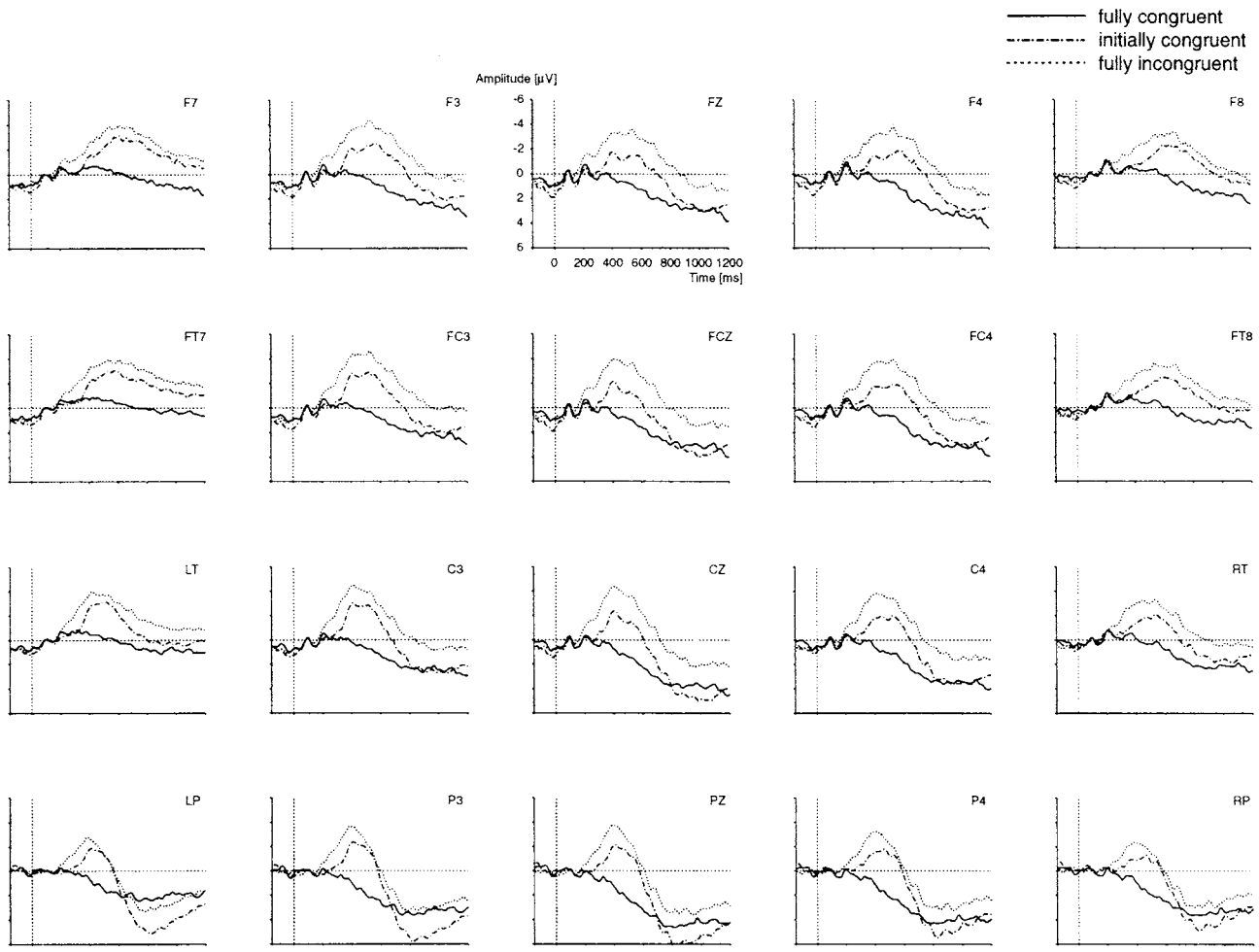


*Figure 1.* Grand average ERPs from 20 scalp sites, to sentence-final words.

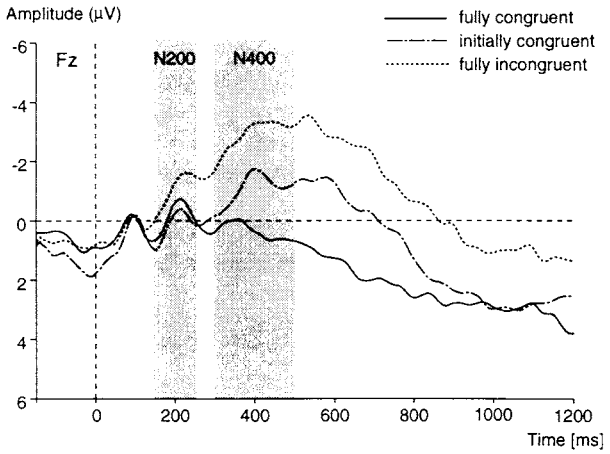


*Figure 2.* Grand average ERPs from the Fz electrode, to sentence-final words.

*Onset latencies.* To establish the onset of the congruity effects, additional analyses were performed. Figure 1 shows that the onset of the effects is most clearly visible over frontal electrodes. Therefore, we decided to analyze this frontal band. Onset latencies were estimated by first separately computing the mean amplitude values for five frontal electrodes in 20 ms latency ranges (bins) that shifted in steps of 10 ms from target onset until 500 ms after target onset (e.g., 0-20, 10-30, etc.). The values for the latency bins were submitted to ANOVAs that tested against the null-hypothesis of zero difference between the a priori selected conditions FC vs. FI and FC vs. IC.

The onset latency analyses for the FC vs. FI comparison revealed that at 150 to 180 ms after word onset, the waveforms of the Fully Incongruent condition momentarily departed from the Fully Congruent condition. After this short-lived congruity effect, the waveforms of the Fully Incongruent condition significantly increased in amplitude approximately 150 ms earlier than the waveform of the Initially Congruent words (i.e., 220 vs. 370 ms after word onset).

*N200 vs. N400.* To establish whether these two congruity effects have statistically distinct scalp distributions, we performed an additional topographical analysis. First, for every subject difference scores between FC and FI were computed for every electrode in the N200 latency window and the N400 latency window. Second, a scaling procedure was performed to avoid that differential amplitude effects between the two latency windows would be incorrectly interpreted as distribution effects. A significant interaction of latency window (N200 and

N400) by electrodes ($F$[28, 560] = 5.50, $MSe$ = 1.76, $p$ < .001) revealed that the congruity effects found in the two latency windows indeed have different spatial distributions across the scalp.

## 4. DISCUSSION

At 150 ms after word onset the ERPs in the Fully Incongruent condition started to diverge from the Fully Congruent and Initially Congruent conditions, indicating that context at that time already had an influence on the spoken-word recognition process. Moreover, we obtained a small early negativity in all conditions, that could be modulated by our experimental conditions (i.e., the N200 was larger in FI than in FC and IC). When taken together with the finding that the early negative shift and the N400 effect have different scalp distributions, this provides evidence in support of the hypothesis that the N200 is a separate negativity preceding the N400. We propose that the early negativity has a functionality that is distinct from the N400 and that reflects a process in word recognition that precedes the integration of a selected word into the sentential context.

We envisage the spoken-word recognition process as follows (cf.[2],[4],[12]): On the basis of an analysis of the initial phoneme(s) of the spoken word a number of lexical candidates are accessed. This is a purely form-driven, bottom-up process. After activation of these candidates top-down context information starts to exert its influence. On the basis of their semantic and syntactic features, candidates in the set are assessed with respect to their goodness of fit within the sentence frame. In the presence of semantic features in the activated set that fit the sentence context well, further incoming acoustic information and top-down contextual information are both used to narrow down the number of candidates to the one that is most compatible with both form and content constraints. This candidate is subsequently integrated in the sentence context. If, however, none of the lexical candidates fit the context well, selection of the proper candidate is difficult and can only be done on the basis of the acoustic information. Once the incongruent word is selected from the set of candidates, integration is attempted.

We propose that the early negativity preceding the N400 reflects the lexical selection process that occurs at the interface of lexical form and contextual meaning (cf. [10]). Analogous to the functional interpretation of the N400, the N200 is indicative of whether the initial assessment of the form-based activated lexical candidates reveals the presence of the semantic features that are required by the contextual specifications. A small N200 is elicited when the set contains semantic features that fit the sentence context (as in FC and IC), a large N200 indicates that the set does not contain semantic features that fit the preceding sentence frame well (as in FI).

## 5. ACKNOWLEDGEMENTS

Note: This manuscript is an extended abstract of a submitted manuscript [13]. Address for correspondence: Daniëlle van den Brink, Max Planck Institute for Psycholinguistics, P.O. Box 310, 6500 AH Nijmegen, The Netherlands. Email: danvdb@mpi.nl.

## 6. REFERENCES

[1] Marslen-Wilson, W.D. (1987). Functional parallelism in spoken word recognition. In Frauenfelder, U.H. & Tyler, L.K. (Eds.), *Spoken word recognition* (pp. 71-102). Cambridge MA: MIT Press.

[2] Marslen-Wilson, W.D. & Welsh, A. (1978). Processing interactions during word recognition in continuous speech. *Cognitive Psychology*, **10**, 29-63.

[3] McClelland, J.L. & Elman, J.L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, **18**, 1-86.

[4] Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, **52**, 189-234.

[5] Kutas, M. & Hillyard, S.A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, **207**, 203-205.

[6] Brown, C.M. & Hagoort, P. (2000). On the electrophysiology of language comprehension: Implications for the human language system. In Crocker, M.W., Pickering, M. & Clifton Jr., C. (Eds.), *Architectures and mechanisms for language processing* (pp. 213-237). Cambridge: Cambridge University Press.

[7] Osterhout, L. & Holcomb, P.J. (1995). Event-related potentials and language comprehension. In Rugg, M.D. & Coles, M.G.H. (Eds.), *Electrophysiology of mind: Event-related brain potentials and cognition* (pp. 171-215). New York: Oxford University Press.

[8] Van Petten, C., Coulson, S., Rubin, S., Plante, E. & Parks, M. (1999). Time course of word identification and semantic integration in spoken language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **25**, 394-417.

[9] Connolly, J.F. & Phillips, N.A. (1994). Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. *Journal of Cognitive Neuroscience*, **6**, 256-266.

[10] Hagoort, P. & Brown, C.M. (in press). ERP effects of listening to speech: Semantic ERP effects. *Neuropsychologia*.

[11] CELEX Dutch database [Electronic database] (1990). Nijmegen, The Netherlands: Centre for Lexical Information [Producer and Distributor].

[12] Zwitserlood, P. (1989). The locus of effects of sentential-semantic context in spoken-word processing. *Cognition*, **32**, 25-64.

[13] Brink, D. van den, Brown, C.M. & Hagoort, P. The N200 component as an electrophysiological indicator of early contextual influences during spoken-word recognition. *Submitted manuscript 2000.*