

ELECTROPHYSIOLOGICAL INSIGHTS INTO LANGUAGE AND SPEECH PROCESSING

P. Hagoort and C.M. Brown

Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

ABSTRACT

Event related brain potentials (ERPs) have been used to study language and speech processing. Two distinct ERP-effects will be discussed: (1) The *N400-effect*. This effect is related to the integration of word meaning into an utterance level representation; (2) The *Syntactic Positive Shift* (SPS). The SPS is related to syntactic processing. Both N400 and SPS were originally observed

in reading, but they can also be observed with speech, with some changes in their latency and distribution.

EVENT RELATED BRAIN POTENTIALS

Cognitive electrophysiology provides a record of various perceptual and cognitive processes as they unfold in real time. The basis for this record are the voltage fluctuations recorded with

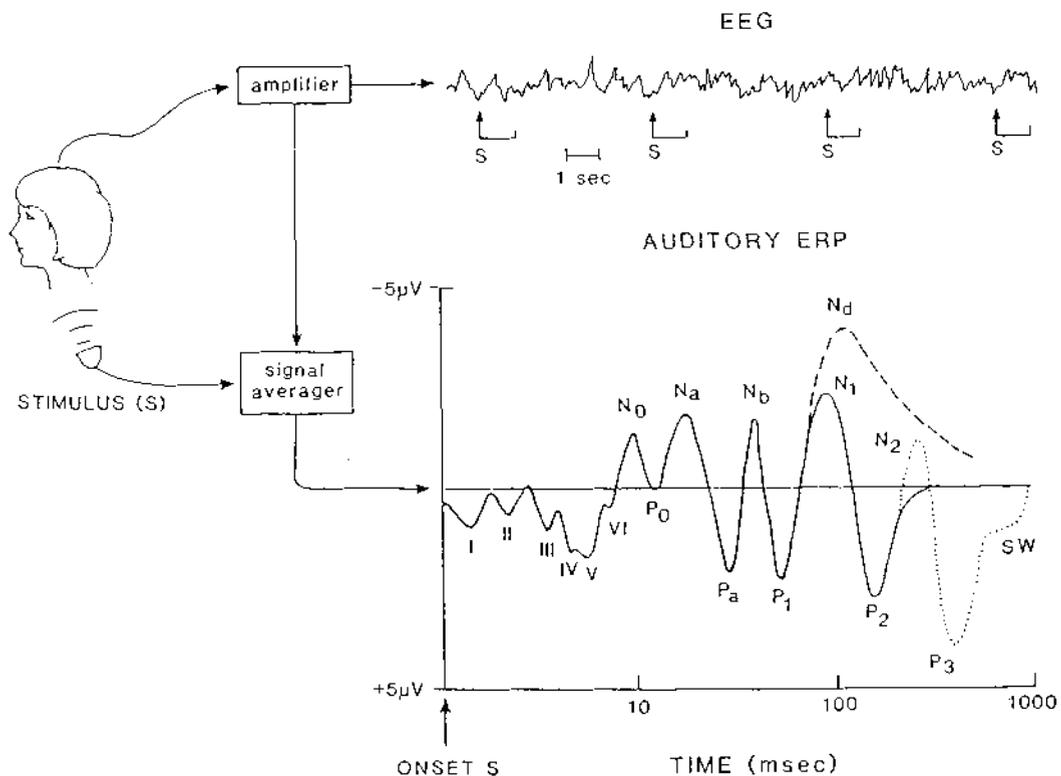


Figure 1 (after [1]): Idealized waveform of a series of ERP components that become visible after averaging the EEG to repeated presentations of a short auditory stimulus. Usually, averaging over a number of stimulus tokens is required to get an adequate signal-to-noise ratio. Along the logarithmic time axis the early brainstem potentials (Waves I-VI), the midlatency components (N_0 , P_0 , N_a , P_a , N_b), the largely exogenous components (P_1 , N_1 , P_2), and the endogenous, cognitive ERP components (N_d , N_2 , P_3 , Slow Wave) are shown. The components with a negative polarity are plotted upwards, the components with a positive polarity are plotted downwards.

the help of electrodes placed on the scalp, known as the electroencephalogram (EEG).

Under the appropriate stimulation conditions, one can derive so-called event related brain potentials (ERPs) from the EEG. Scalp-recorded ERPs reflect the summation of synchronous post-synaptic activity of a large number of neurons. ERPs differ from background EEG in that they reflect brain electrical activity time-locked to particular stimulus events. Establishing a reliable ERP trace normally requires averaging over a series of ERP recordings to tokens of the same stimulus type. The resulting average waveform typically includes a number of positive and negative peaks, often referred to as ERP-components (see Figure 1). Usually, the peaks in the ERP waveform are labelled according to their polarity (*N* for negative, *P* for positive) and their average latency in milliseconds relative to the onset of stimulus presentation (e.g., *N400*, *P300*). In some cases, the ERP peaks get a functionally defined label (*SPS* for syntactic positive shift; *ERN* for error-related negativity). ERPs are recorded from a number of electrodes distributed over the scalp. Often they have a characteristic distribution, showing larger amplitudes at some sites than at others. These distributional characteristics can be helpful in identifying a certain component.

For the purposes of psycholinguistically oriented ERP research, the most informative ERPs belong to the class of so-called "endogenous" components. Endogenous components are relatively insensitive to variations in physical stimulus parameters (e.g., size, intensity), but highly responsive to the cognitive processing consequences of the stimulus events. The modulations in amplitude or latency of an endogenous ERP as a consequence of some experimental manipulation, usually form

the basis for making inferences about the nature of the underlying cognitive processing events.

For research on language and speech processing, particular two characteristics of ERPs are of relevance. The first is the *multidimensional* nature of the ERP waveform. ERPs can vary along a number of dimensions: specifically, the latency at which an ERP component occurs relative to stimulus onset, its polarity, its amplitude, and its amplitude distribution over the recording sites. On the basis of these characteristics it is reasonable to assume that different types of ERP peaks (e.g. positive peaks vs. negative peaks) are generated by, at least in part, non-overlapping neuronal populations. Insofar as the involvement of different neuronal ensembles implies qualitatively different processing events, in principle these processing events can show up as qualitatively different in the overall ERP waveform. This characteristic makes ERPs a useful addition to the recording of unidimensional measures, such as reaction times. For instance, if in sentence processing the electrophysiological signatures of semantic integration processes and parsing operations turn out to be qualitatively different, ERPs might provide us with a crucial tool for testing how and at what moments in time the process of assigning a structure to the incoming string of words, and interpreting this string semantically, influence each other.

The second important characteristic of ERPs is that they provide a *continuous, real-time* measure. This high temporal resolution of ERPs is unmatched by other brain imaging techniques such as PET and fMRI. Like speeded reaction-time (RT) measures in the more classical psycholinguistic tasks, such as naming, lexical decision, and word or phoneme monitoring, ERPs are tightly linked to the temporal organization of ongoing language processing events. But in

contrast to RT measures, ERPs provide a continuous record throughout the total processing epoch and beyond. Therefore, it is possible to monitor not only the immediate consequences of a particular experimental manipulation (e.g., a syntactic or semantic violation), but also its processing consequences further downstream. This feature enabled us to show that the impossibility of assigning the preferred structure to an incoming string of words has consequences for lexical-semantic integration processes further downstream in the sentence (see below) [2].

N400-EFFECTS

The N400 was first reported in a paper by Kutas & Hillyard (1980) [3]. These authors presented subjects with a variety of sentences either ending in a word that was semantically congruous with the sentence context (e.g., "He shaved off his mustache and beard") or ending in a semantic anomaly ("I take coffee with cream and dog"). The semantically anomalous words elicited a negative component with a centro-parietal maximum on the scalp, and a latency that peaked around 400 ms. This component has since become known as the N400, and the difference between the N400 amplitude in the experimental and the control conditions has become known as the N400 effect.

Since its discovery, it has become clear that N400 effects are not elicited by only semantic *violations*. This can be illustrated by the following result from one of our studies [4]. We presented subjects with sentences that were identical, with the exception of a highly expected word in sentence-medial position (e.g., "Jenny put the sweet in her *mouth* after the lesson) versus a word that made perfect sense but was less expected in this position (e.g., "Jenny put the sweet in her *pocket* after the lesson"). Figure 2 shows the ERP waveforms to the more and less

expected words, preceded and followed by one word. As can be seen in this figure, the N400 to the less expected word 'pocket' is larger than to the word 'mouth' which is the more expected continuation of the context. This probably reflects the different degree to which these words can be readily integrated within the higher order representation of their preceding sentential-semantic context.

Across many N400 studies, the following general characteristics are known to hold for the N400: (a) Each open-class word elicits an N400. (b) The amplitude of the N400 is inversely related to the cloze probability of a word in sentence context. The better the semantic fit between a word and its context, the more reduced the amplitude of the N400. (c) The amplitude of the N400 varies with word position, such that the first content word in a sentence produces a larger negativity than content words in later positions. This amplitude reduction is most likely due to the increasing semantic constraints throughout the sentence. (d) N400 effects are obtained in sign language, but not with violations of contextual constraints in music.

Importantly, N400 effects are not only observed with visual language input, but also with speech input. The most important difference between N400 effects to written and spoken words is their onset latency. Whereas the N400 effect in the visual modality usually onsets at about 250 ms, with spoken input the onset can be up to 200 ms earlier. This means that on average, N400 effects to speech start to emerge well before the end of a word.

Recent research suggests that the amplitude of the N400 is related to lexical-semantic integration processes [5]. That is, once a word has been accessed in the mental lexicon, its meaning has to be integrated into an overall representation of the current

HIGH CLOZE vs. LOW CLOZE PROBABILITY WORDS

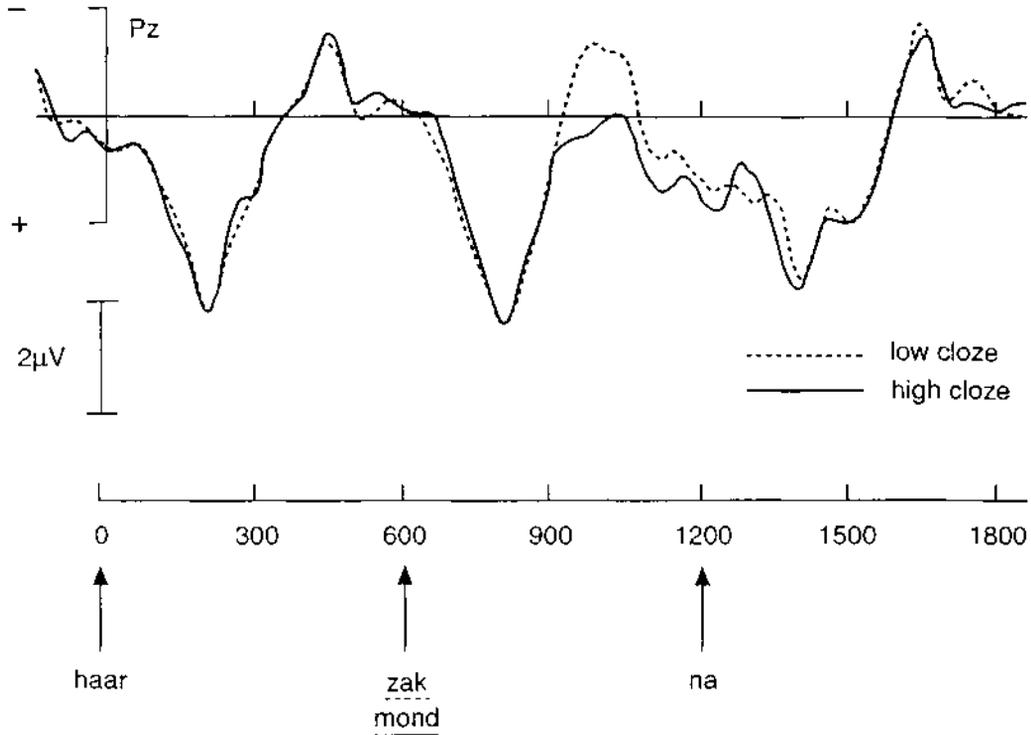


Figure 2: Grand-average waveform for electrode site Pz, for sentence-medial words with a high cloze probability and a low cloze probability. Sentences were presented word by word on the center of a computer screen at a rate of one word per 600 ms. The cloze target is preceded and followed by one word. The translation of the Dutch example sentence is "Jenny put the sweet in her pocket/mouth after the lesson." The waveforms represent the part of the sentence that is underlined.

word or sentence context. The easier this integration process is, the smaller the amplitude of the N400 becomes.

The early onset of N400 effects in speech attests to the immediacy of lexical-semantic integration processes in this modality.

THE SYNTACTIC POSITIVE SHIFT

In recent years a number of ERP studies on syntactic processing have clearly shown that the ERP responses to violations of syntactic preferences are qualitatively different from the classical N400 effect [2][6].

In one of our studies, we had subjects

read sentences that violated the agreement between the subject nounphrase and the finite verb, as in the following example sentences (literal translation in English between brackets; the word that renders the sentence ungrammatical (the Critical Word [CW] and its counterpart are italicized):

"Het verwende kind *gooit* het speelgoed op de grond."

(The spoiled child *throws* the toys on the floor.)

* "Het verwende kind *gooien* het speelgoed op de grond."

(The spoiled child *throw* the toys on the floor.)

AGREEMENT CONDITION, Electrode Pz

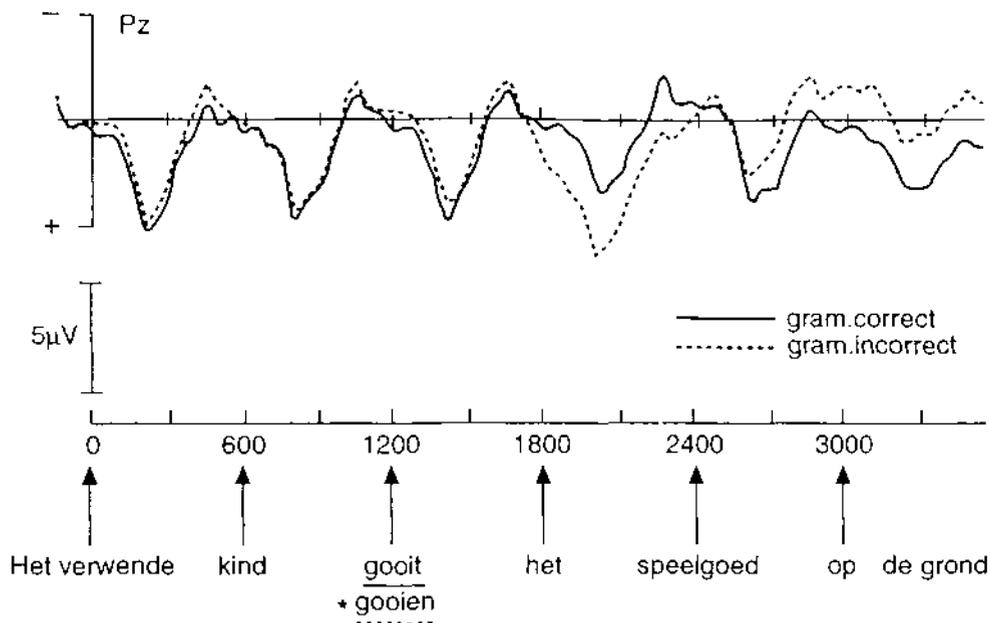


Figure 3: Grand-average waveform for electrode site Pz, for the grammatically correct and incorrect Critical Words (CW). The CW is preceded by two and followed by three words. Sentences were presented word by word on the center of a computer screen at a rate of one word per 600 ms. The translation of the example sentence is "The spoilt child throws/throw the toy on the ground."

The basic pattern of results that we observed is shown in Figure 3 for a posterior midline electrode (Pz). The CW is preceded by two words and followed by three words.

As can be seen, the ERP waveform to the incorrect CW shows a positive shift in comparison with its correct counterpart. This positive shift is widely distributed over the recording sites and has a centro-parietal maximum. Based on its sensitivity to syntactic aspects of a sentence, we have labeled this effect the SPS (i.e., Syntactic Positive Shift). The onset of the SPS is at about 500 ms after presentation of the incorrect CW. A similar pattern of results is obtained for a number of other syntactic violations in both Dutch and English.

As can be seen, the SPS is replaced by a negative shift on word positions following the CW. These are N400 effects, indicating the increased

difficulty of integrating words into the sentence context following a syntactic violation.

As holds for the N400, the SPS is not elicited by only syntactic violations. In general, an SPS can be observed when a syntactic preference can no longer be maintained. That is, the word in the sentence that renders the preferred syntactic structure impossible, elicits an SPS. An example in case are so-called syntactically ambiguous sentences. Very often part of a sentence can be assigned more than one syntactic structure. For instance, in the utterance "The pope greets the priest and the monk...", the noun 'monk' can go together with 'priest' to form the object of the sentences (e.g., "The pope greets the priest and the monk at the annual meeting"). Alternatively it can start a new clause (e.g., "The pope greets the priest and the

monk welcomes the cardinal"). For reasons of processing economy, the first (conjoined-NP) reading is preferred over the second (Sentence conjunction) reading. The verb that renders the preferred syntactic assignment impossible ('welcomes' in the example), therefore, elicits an SPS. In short, the SPS seems to signal that the initially assigned syntactic structure can no longer be maintained, and that some form of reanalysis has to be initiated.

To date, only two studies have tested for the occurrence of an SPS to syntactic violations in spoken sentences. Osterhout and Holcomb [7] report a somewhat earlier onset of the SPS during the perception of continuous speech. In our own study, however, the onset of the SPS in continuous speech was quite similar to that in the visual modality. Although more research needs to be done with continuous speech input, current results tentatively suggest that the signal for reanalysis is relatively insensitive to the rate at which words are presented.

CONCLUSIONS

From these results the following general conclusions can be drawn:

(1) Electrophysiological recordings provide a real-time neurophysiological measure of language and speech processing with a temporal resolution that is far superior in comparison with other brain imaging techniques such as PET and fMRI. These latter methods, however, have a much better spatial resolution than ERPs. For a full understanding of the neurobiological basis of language and speech, we have to rely on the combined use of different brain imaging techniques.

(2) The existence of different ERP responses to aspects of semantic and syntactic processing suggests different underlying brain states for semantic

integration and parsing. To the degree to which the SPS and the N400 individuate different sets of neural generators, and to the degree to which these different sets of neural generators (directly or indirectly) correspond to different cognitive states, it can be concluded that the processing mechanism for the computation of syntactic structures is different from that for the computation of the meaning of an utterance. In other words, the brain honours the distinction between syntax and semantics.

(3) The observed ERP effects are independent of modality. That is, both with written and spoken input N400 effects and SPS are observed. However, the effects seem to be earlier in continuous speech, especially for the N400. This attests to the speed at which speech has to be processed.

(4) The findings of qualitatively different neurophysiological responses to semantic and syntactic processing, suggests that in further research additional ERP effects sensitive to, for instance, early phonological processing might be obtained. Some recent findings are suggestive of this possibility [8][9].

ACKNOWLEDGEMENT

The research reported in this paper was supported by a grant from the Dutch Science Foundation (NWO), with grant number 400-56-384.

REFERENCES

- [1] Hillyard, S.A., & Kutas, M. (1983), "Electrophysiology of cognitive processing." *Ann. Rev. Psychol.*, vol. 34, pp. 33-61.
- [2] Hagoort, P., Brown, C.M., & Groothusen, J. (1993), "The syntactic positive shift (SPS) as an ERP-measure of syntactic processing." *Language and Cognitive Processes*, vol. 8, pp. 439-483.
- [3] Kutas, M., & Hillyard, S.A. (1980), "Reading senseless sentences: Brain

- potentials reflect semantic incongruity." *Science*, vol. 207, pp. 203-205.
- [4] Hagoort, P., & Brown, C.M. (1994), "Brain responses to lexical ambiguity resolution and parsing" In: Ch. Clifton Jr., L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing*. Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 45-80.
- [5] Brown, C.M., & Hagoort, P. (1993), "The processing nature of the N400: Evidence from masked priming." *Journal of Cognitive Neuroscience*, vol. 5, pp. 34-44.
- [6] Osterhout, L., & Holcomb, P.J. (1992), "Event-related brain potentials elicited by syntactic anomaly." *Journal of Memory and Language*, vol. 31, pp. 785-806.
- [7] Osterhout, L., & Holcomb, P.J. (1993), "Event-related potentials and syntactic anomaly: Evidence of anomaly detection during the perception of continuous speech." *Language and Cognitive Processes*, vol. 8, pp. 413-437.
- [8] O. Conolly J.F., & Phillips, N.A. (1994), "Event-related potential components reflect phonological and semantic processing of the terminal words of spoken sentences." *Journal of Cognitive Neuroscience*, vol. 6, pp. 256-266.
- [9] Praamstra, P., Meyer, A.S., Levelt, W.J.M. (1994), "Neurophysiological manifestations of phonological processing: Latency variation of a negative ERP component time-locked to phonological mismatch." *Journal of Cognitive Neuroscience*, vol. 6, pp. 204-219.