



ERP effects of listening to speech compared to reading: the P600/SPS to syntactic violations in spoken sentences and rapid serial visual presentation

Peter Hagoort*, Colin M. Brown

Max Planck Institute for Psycholinguistics, Wundtlaan 1, NL-6525 XD Nijmegen, The Netherlands

Received 29 January 1997; received in revised form 10 January 2000; accepted 11 February 2000

Abstract

In this study, event-related brain potential effects of speech processing are obtained and compared to similar effects in sentence reading. In two experiments sentences were presented that contained three different types of grammatical violations. In one experiment sentences were presented word by word at a rate of four words per second. The grammatical violations elicited a Syntactic Positive Shift (P600/SPS), 500 ms after the onset of the word that rendered the sentence ungrammatical. The P600/SPS consisted of two phases, an early phase with a relatively equal anterior–posterior distribution and a later phase with a strong posterior distribution. We interpret the first phase as an indication of structural integration complexity, and the second phase as an indication of failing parsing operations and/or an attempt at reanalysis. In the second experiment the same syntactic violations were presented in sentences spoken at a normal rate and with normal intonation. These violations elicited a P600/SPS with the same onset as was observed for the reading of these sentences. In addition two of the three violations showed a preceding frontal negativity, most clearly over the left hemisphere. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: P600/SPS; LAN; Natural speech; RSVP; Syntactic processing

1. Introduction

In this study, we report Event Related Potential (ERP) effects to spoken and written sentences. Our aim here is to determine some of the commonalities and differences in electrophysiological responses related to syntactic processing via the auditory and visual modality. Especially the time course and the distribution of the ERP effects for both kinds of input not only provide further evidence about their dependence or independence on input modality, but this between-modality comparison can also be helpful in further delineating the functional nature of syntax-related ERP effects.

Over the last 15 years or so, distinct ERP effects

have been related to two separate aspects of sentence reading (see [12], for an overview). One is the so-called N400, related to semantic processing [17], see Hagoort and Brown, this issue). Other ERP effects have been reported in relation to syntactic processing. These include the LAN [7,20] and the P600/SPS (for a review, see [12]).

Although these qualitatively distinct ERP effects indicate that the brain honours the distinction between semantic and syntactic processing, this does not allow the inference that any of these effects are language specific. In fact, there is insufficient evidence with respect to all of the language-relevant ERP componentry for the claim of language specificity. Presumably, (a subset of) the generators of these ERP effects are sensitive to other cognitive processes as well (cf. [8,29]). However, within the context of language processing these ERP effects are differentially affected by

* Corresponding author. Fax: +31-24-3521213.
E-mail address: peha@mpi.nl (P. Hagoort).

distinct types of linguistic information. This is the crucial requirement for fruitfully exploiting language-relevant ERP effects for studies on the neurocognitive architecture of the human language system.

Among the language-related ERP effects one well-established effect is linked to structural aspects of the sentence. This effect has been variously labeled the P600 [23] or the SPS (Syntactic Positive Shift; [11]). This ERP effect, here referred to as P600/SPS, can be elicited by words that render a sentence ungrammatical. For visually presented sentences, the onset of this effect is usually in the order of 500 ms.

Both N400 effects and the P600/SPS can also be obtained for more subtle aspects of semantic and syntactic processing than straightforward violations of semantic and syntactic constraints (see for recent overviews, [12,18,25]). However, in this study we exploited the violation paradigm to obtain P600/SPS effects in natural speech, for direct comparison with their classical visual counterparts.

In the first experiment we investigated the characteristics of the P600/SPS when subjects were reading four words per second, a rate that comes close to listening to speech. In the second experiment we investigated the characteristics of the P600/SPS with speech as input.

The P600/SPS has been observed in response to a number of different types of syntactic violations in English, Dutch, and German [7,11,21,23]. For example, Hagoort et al. [11] compared ERPs to Dutch sentences that violated the agreement between the subject noun phrase and the finite verb (e.g. **The spoiled child throw the toys on the floor.*'). They observed a positive shift starting at about 500 ms post-stimulus onset in response to the finite verb (underlined in the example) that rendered the sentence ungrammatical. The same effect was also obtained following violations of normal word order within Dutch noun phrases (e.g. **the expensive very tulip*'). Likewise, Osterhout and Holcomb [23] reported similar ERP effects to violations of verb subcategorization and phrase structure constraints in English. In one of their experiments subjects were asked to read active sentences in which the finite verb was followed by a reduced relative clause and a verb phrase (e.g. *The broker persuaded to sell the stock was sent to jail*'). In some cases, however, the sentences were ungrammatical because an intransitive verb was used which does not allow a reduced relative reading (e.g. **The broker hoped to sell the stock was sent to jail.*'). These sentences became ungrammatical at the auxiliary verb *was*. In the ungrammatical sentences this verb elicited a positive shift starting 400–500 ms after stimulus onset and lasting in the order of 400 ms.

The dissociation of N400 and P600/SPS effects is further substantiated by the finding that a P600/SPS

can also be obtained to syntactic violations in sentences in which the usual semantic constraints no longer apply [10]. For instance, the agreement violation in semantically odd sentences (e.g. *The boiled watering-can smoke the telephone in the cat.*') results in a similar P600/SPS as the agreement violation in normal sentences (e.g. **The spoiled child throw the toys on the floor.*'). Furthermore, a straightforward syntactic violation is not required to elicit a P600/SPS. It has been shown that the P600/SPS is also elicited to a word in the sentence that renders the preferred structural assignment among the alternative assignment options impossible [1,12,27]. Although there is no evidence that the P600/SPS is a language-specific ERP effect, in the context of sentence processing it clearly supports the view that syntactic processing is domain specific (cf. [4]).

So far, almost all studies reporting P600/SPS effects have used a relatively slow visual presentation, in which words were presented sequentially at a rate of about 600 ms per word. Compared to the rate of normal reading (about four words per second) and to the rate of connected speech, these presentation parameters do not match the on-line characteristics of normal sentence processing. Thus, although the conditions resulting in N400 and P600/SPS effects suggest that the brain honours the distinction between semantic and syntactic sentence-level processing, with respect to the P600/SPS not much is known about its characteristics when a normal reading rate is imposed on the subjects, or when the input is speech. This information is, however, critical if one wants to relate ERP effects to aspects of on-line sentence processing in reading and listening.

To date, we are not aware of any study testing the P600/SPS with a visual presentation rate that approaches normal reading speed, that is, in the order of four words per second. Until now five studies have investigated the P600/SPS with connected speech as input [6,7,13,24,29]. In these studies clear P600/SPS effects were observed to the syntactic anomalies. However, especially in the Osterhout and Holcomb [24] study there were two differences with the P600/SPS observed in the visual modality. First, the effect had a slightly different topographic distribution. In connected speech, the P600/SPS was more restricted to posterior sites and more widely distributed over the right than the left hemisphere compared to the effects in the visual modality. Second, the latency seemed to be earlier for the spoken than for the written sentences. This earlier latency in speech was also observed in the study by Patel et al. [29]. However, this finding was not corroborated in the Friederici et al. and the Hahne and Friederici studies [6,13], in which similar onset latencies of the P600/SPS were observed with speech as for written input.

In addition, the syntactic anomalies in connected speech elicited an early negative effect that was most clear over (left) anterior sites [7,13,24,]. However in the Hahne and Friederici study [13] this effect is much earlier than in the studies by Osterhout and Holcomb [24] and Friederici et al. [7], where the effect was found around 400 ms after word onset. Hahne and Friederici [13] found a frontal negativity between 100 and 300 ms, which they refer to as ELAN.

At present it is still unclear whether these differences between the results of studies with relatively slow visual presentation and the studies with connected speech are actually due to differences in modality, or relate to differences in the rate of presentation. We therefore decided to run two experiments, one in which the rate of presentation in the visual modality was close to normal reading, and the other with connected speech. In this way we hoped to disentangle the rate and modality contributions to possible differences between a visual and an auditory P600/SPS.

An additional reason for our interest in testing the P600/SPS at a higher visual presentation rate and with connected speech evolved from an unexpected outcome of our earlier studies [10,11]. In these studies there was one syntactic violation that neither elicited a P600/SPS nor any other ERP effect to the word that rendered the sentence ungrammatical. This was a violation of the subcategorization constraint, according to which an intransitive verb cannot be followed by an object noun phrase (e.g. **The son of the rich industrialist boasts the car of his father.*). Compared to the correct counterpart of these sentences (e.g. *The son of the rich industrialist borrows the car of his father.*), no ERP effect was seen to the critical noun in object position. We have proposed that in this violation syntactic and semantic constraints are intertwined. As a result, the violation might elicit temporally overlapping P600/SPS and N400 effects, leading to a cancellation of each of these effects in the overall ERP waveform [10]. However, it is not necessarily the case that the time course of P600/SPS and N400 effects is similarly affected by rate or modality changes. Thus we were interested to see whether for the subcategorization violation an ERP effect would emerge under different presentation parameters.

¹ In Dutch it is a legal but infrequent option to have a nominalized infinitive in subject position. In sentence 2, the sequence *“Het verwende kind gooien”* can be read as a nominalized infinitival cluster in grammatical subject position. In this reading, only the word following the verb renders the sentence ungrammatical. However, it is clearly the preferred reading to interpret the verb as a finite verb marked for plural, and this was substantiated by the results of online grammaticality judgement pretests on the materials.

2. Materials

The same materials were used as in the Hagoort et al. study [11]. These materials consisted of a set of 360 Dutch sentences, half of which are grammatically correct, the other half with a grammatical violation. Apart from the syntactic violation, the 180 ungrammatical sentences were identical to the 180 grammatical sentences. The grammatical violations were of three different types: (1) violation of number agreement between the subject noun phrase and the finite verb, (2) violation of verb subcategorization, and (3) violation of phrase structure.

The agreement violations all consisted of number violations between verbs and nouns within SVO (Subject Verb Object) and VSO (Verb Subject Object) sentences. Violations occurred on either finite verbs or subject nouns within the same sentence. The following example gives both the grammatically correct and incorrect versions of an SVO agreement violation (literal translations in English between brackets; the word that renders the assignment of a grammatical or preferred sentence structure impossible (i.e. the critical word, CW) and its correct counterpart are underlined).¹

1. *Het verwende kind gooit het speelgoed op de grond.* (The spoilt child throws the toys on the floor.)
2. **Het verwende kind gooien het speelgoed op de grond.* (The spoilt child throw the toys on the floor.)

The subcategorization violations involved obligatory intransitive verbs. Such verbs cannot take a noun as direct object. This constraint was violated in the ungrammatical versions of the sentences as is illustrated in the following example.

3. *De zoon van de rijke industrieel leent de auto van zijn vader.* (The son of the rich industrialist borrows the car of his father.)
4. **De zoon van de rijke industrieel pocht de auto van zijn vader.* (The son of the rich industrialist boasts the car of his father.)

The phrase structure violations all consisted of nouns preceded by permutations of adverbs and adjectives. In Dutch, it is a violation of phrase structure constraints to have an adjective-adverb-noun sequence. For example:

5. *De echtgenoot schrikt van de nogal emotionele reactie van zijn vrouw.* (The husband [is startled] by the rather emotional response of his wife.)
6. **De echtgenoot schrikt van de emotionele nogal reactie van zijn vrouw.* (The husband [is startled] by the emotional rather response of his wife.)

In this example, as in all of the phrase-structure violation sentences, the actual violation occurs on the noun following the critical word (i.e. on *reactie*). It is only at this point that the sentence can no longer be continued in a grammatically correct manner, since the adjective-adverb combination could be part of a more complex adjective-adverb-adjective-noun sequence (e.g. *'the emotional rather violent response'*). However, in the ungrammatical sentences the P600/SPS was already elicited by the adverb [10,11]. We have interpreted this as indicating that parsing difficulties already arise at the adverb, since it signals that a more complex, less frequent structure has to be assigned than the preferred adjective-noun sequence (e.g. *'emotional response'*).

The 360 sentences were distributed over two experimental lists of 180 sentences each. The separate lists contained 90 correct and 90 incorrect sentences, with 30 correct and 30 incorrect sentences of each violation type. The members of a pair of incorrect and correct companion sentences were assigned to different lists. Each list was preceded by a set of 30 practice sentences (15 grammatically correct and 15 incorrect). Each list was presented to a different group of subjects.

Each list of experimental items was divided into three blocks. After the practice session, and after the first and second block, subjects were given a pause. Each block of experimental sentences was preceded by four filler sentences (for additional details of the materials and the design, see [11]).

3. Experiment 1: rapid serial visual presentation (RSVP)

3.1. Method

3.1.1. Subjects

All subjects in this experiment had normal or corrected-to-normal vision. None of the subjects had any neurological impairment or had experienced any neurological trauma according to their responses on a questionnaire. The handedness of the subjects was established via an adapted Dutch version of the Oldfield Handedness Inventory [22]. Subjects were paid for their participation. A total of 32 subjects (14 female; 32 right-handers) participated in the experiment, 16 on each list.

3.1.2. Procedure

The sentences were displayed word by word in white lower-case letters (the first word of each sentence began with a capital letter) against a dark grey background, in the centre of a high-resolution computer screen. Viewing distance was between 70 and 80 cm. The words subtended a vertical visual angle of approximately 3°. Each word was presented for 129 ms,

followed by a blank screen for another 129 ms. Then the next word of the sentence appeared on the screen, etc. The length of the inter-trial interval (ITI) was variable, and depended on the number of words in each sentence.

The subjects were instructed to focus on comprehending the whole sentence. No additional task demands were imposed. The subjects were told that some sentences would be grammatically incorrect, but they were given no information concerning the kinds of grammatical errors that would occur. The entire session, including electrode application and removal lasted at most 2 hours.

3.1.3. EEG-recording

The EEG was recorded from 7 tin electrodes in an electrode cap. Each electrode was referred to the left mastoid. Three electrodes were placed according to the international 10–20 system [14] over midline sites at Fz, Cz, and Pz locations. Four electrodes were placed laterally over symmetrical anterior and posterior positions. The anterior temporal electrodes (ATL, ATR) were located one half of the distance between F7 and T3, and F8 and T4 respectively. Symmetrical posterior temporal electrodes (PTL, PTR) were placed lateral, by 30% of the interaural distance, and 13% posterior to the vertex. Vertical eye movements and blinks were monitored via a supra- to sub-orbital bipolar montage. A right to left canthal bipolar montage was used to monitor for horizontal eye-movements. Activity over the right mastoid bone was recorded on an additional channel to determine if there were differential contributions of the experimental variables to the two presumably neutral mastoid sites. No such differential effects were observed.

The EEG and EOG recordings were amplified with Nihon Kohden AB-601G bio-electric amplifiers, using a Hi-Cut of 30 Hz and a time constant of 8 s. Impedances were kept below 5 k Ω . The EEG and EOG were digitized on-line with a sampling frequency of 200 Hz. Sampling started 150 ms before the presentation of the sentences and continued for a period of 5.5 s. The experimental trials were stored along with condition codes for off-line averaging and data analysis.

3.2. Results

For all subjects, after removal of artifact-contaminated trials (3%), average waveforms were computed for the correct and incorrect sentences of the three violation types separately. Figs. 1, 3, and 5 show the results for the agreement violation, the subcategorization violation and the phrase structure violation, respectively. The waveforms span the presentation epoch of the critical word and four following words. As the

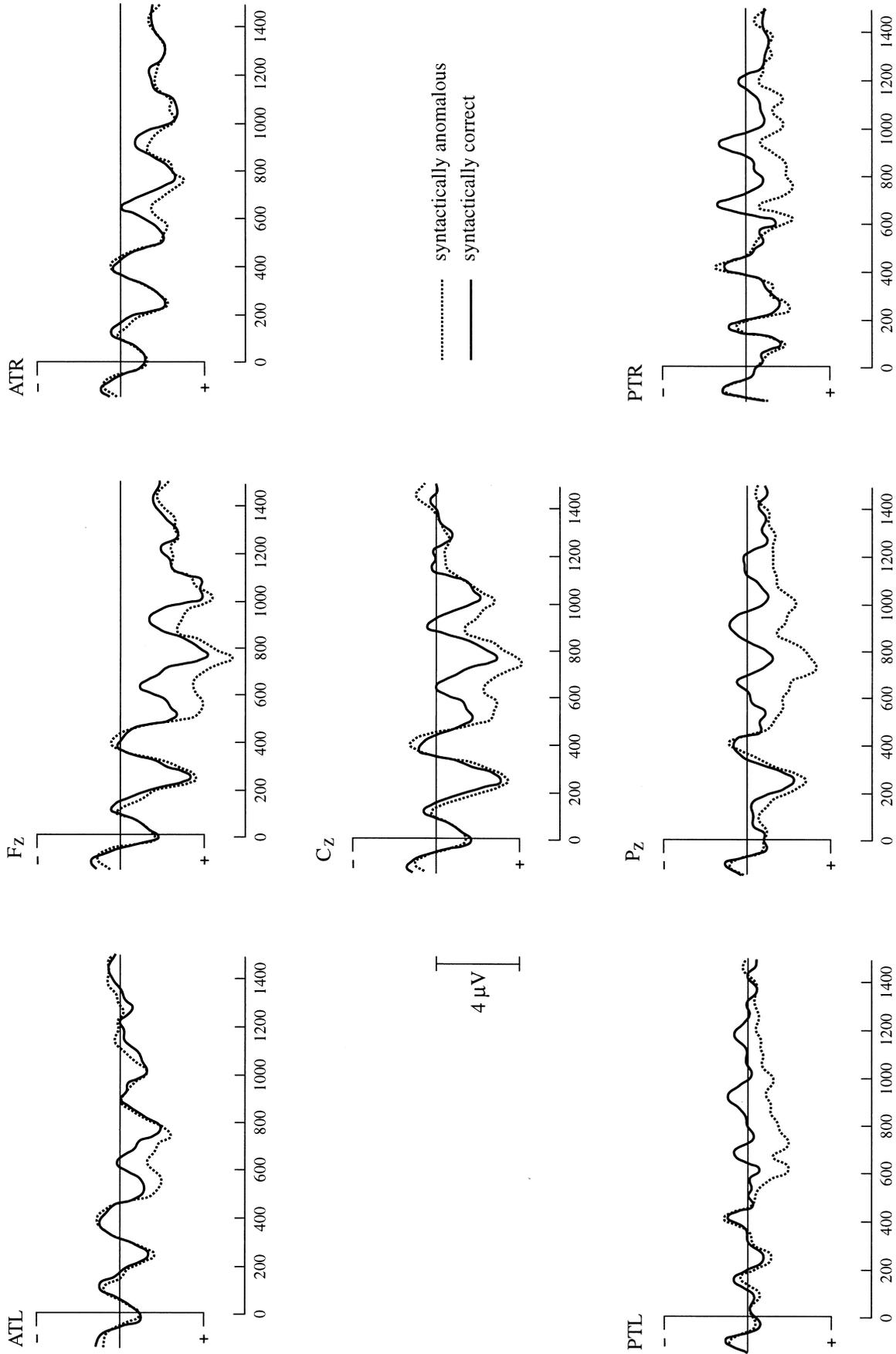


Fig. 1. Agreement condition, RSVP: Grand average ERPs (across all subjects and items), from 7 scalp sites, for the grammatically correct and incorrect critical words (CW). The onset of the CW is at zero milliseconds. The waveforms display the ERPs to the CW and the four following words. The time axis is in milliseconds. Each plot represents approximately 931 trials. Negativity is plotted upwards in this and all following figures.

figures show, to a greater or lesser extent for all violation types a positive shift occurs at about 500 ms. Especially at posterior sites this positive shift is fairly long-lasting. Both on the basis of results of earlier studies [10,11] and visual inspection of the waveforms, we decided to do statistical analyses in two latency windows: One from 500–750 ms after presentation of the critical word, and one in the latency range of 750–1000 ms.

In addition, especially in the subcategorization condition, the P600/SPS seemed to be preceded by a negative shift between 350 and 450 ms. Statistical analyses were therefore also done for this latency window.

Repeated measures ANOVAs were done separately for the three violation types. The analyses were based on the mean amplitude in the three windows specified above, relative to a 150 ms baseline preceding the critical word. The relevant factors in the ANOVAs were grammaticality (grammatical, ungrammatical) and site (ATL, Fz, ATR, Cz, PTL, Pz, PTR). The Greenhouse–Geisser correction was applied when evaluating effects with more than one degree of freedom in the numerator ([9]; see [32]). The adjusted degrees of free-

dom and *P*-values will be reported. Effects involving the factor site are only reported for significant interactions with grammaticality. We also explicitly tested for the topographical differences between the first and second P600/SPS latency window. This was done after rescaling the effect sizes in the two windows as *z*-scores such that for each window means and variances across electrode sites became 0 and 1, respectively [31]. This procedure is equivalent to the procedure suggested by McCarthy and Wood [19].

3.2.1. Agreement

As can be seen in Fig. 1, waveforms start to diverge at about 500 ms after the critical word, with a positive shift for the ungrammatical condition. This positive shift has an initial phase in which it is roughly equally distributed over anterior and posterior sites. This early phase is followed by a later one with a more posterior distribution and with a longer duration over posterior than anterior sites. Fig. 2 shows the topographic distribution of the first and second 250 ms of the syntactic positive shift (i.e. from 500 to 750 ms, and from 750 to

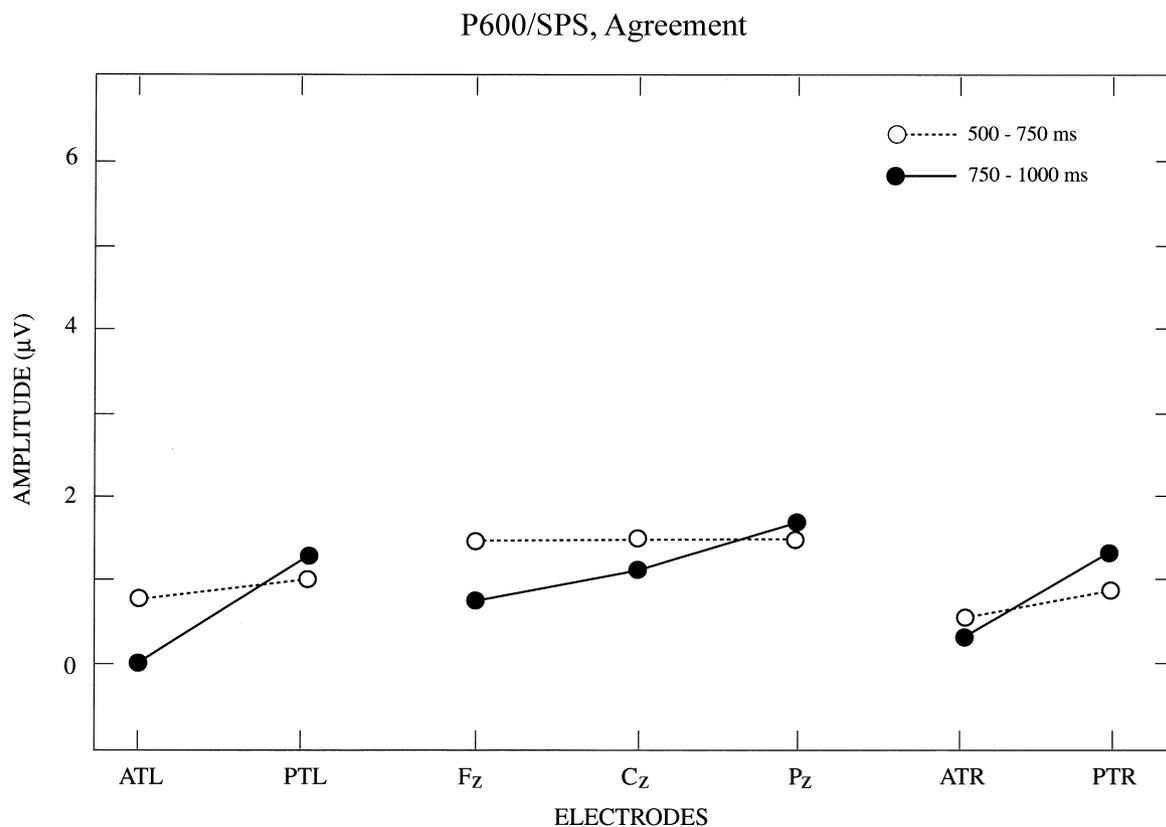


Fig. 2. Agreement condition, RSVP. The topographical distribution of the P600/SPS in the 500–750 ms latency window and in the 750–1000 ms latency window following presentation of the critical word (CW). For these two latency windows, the mean amplitude differences between the syntactically correct and the syntactically anomalous CW are displayed for two left hemisphere sites (ATL, PTL), three midline sites (Fz, Cz, Pz), and two right hemisphere sites (ATR, PTR).

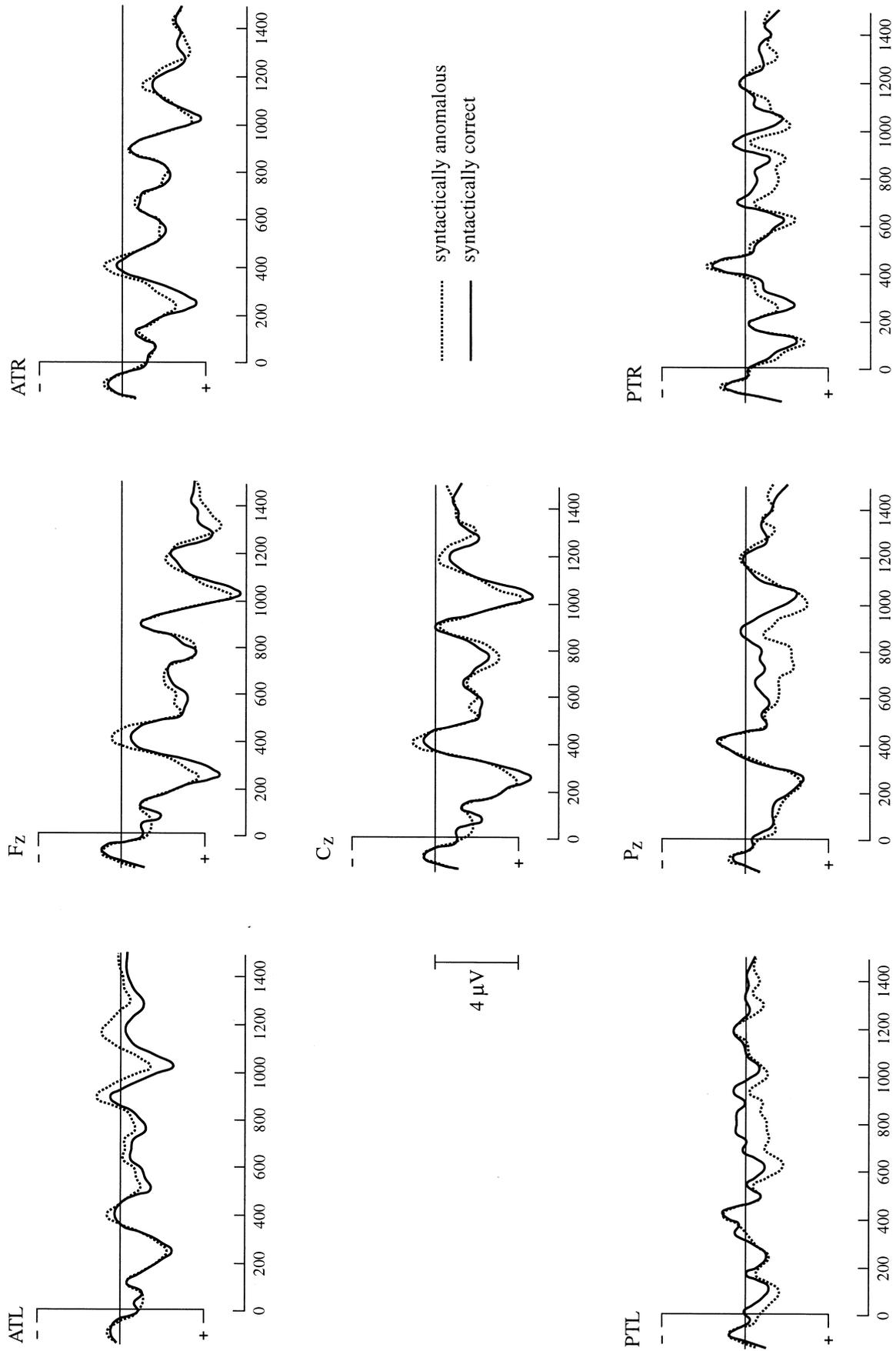


Fig. 3. Subcategorization condition, RSVF: Grand average ERPs (across all subjects and items), from 7 scalp sites, for the grammatically correct and incorrect critical words (CW). The onset of the CW is at zero milliseconds. The waveforms display the ERPs to the CW and the four following words. Each plot represents approximately 931 trials.

1000 ms). On the whole, the effect is largely symmetric over both hemispheres.

The ANOVAs showed significant overall effects of grammaticality for both the 500–750 ms window ($F(1,31)=26.91$, $MSe=7.36$, $P < 0.0001$) and the 750–1000 ms window ($F(1,31)=17.86$, $MSe=7.53$, $P < 0.0005$). In addition, for the second window but not for the first, a significant interaction with site was obtained ($F(1,31)=4.89$, $MSe=1.60$, $P < 0.05$). Separate analyses for individual electrodes showed significant effects for all electrodes in the 500–750 ms window, and the same for the 750–1000 ms window with the exception of the ATL and ATR electrodes. An ANOVA in which the topographical differences were directly tested after z -score normalization resulted in a significant window by electrode interaction ($F(1,31)=5.37$, $MSe=0.32$, $P < 0.05$). The statistical analyses thus support the conclusion that for the first window an equal distribution over the scalp was obtained, with a more posterior distribution for the second window. The small negative shifts preceding the P600/SPS at Fz and Cz failed to reach significance ($F_s < 1$).

3.2.2. Subcategorization

In contrast to previous studies [10,11], in this study grammaticality effects seem to be present. However, as Fig. 3 shows, the P600/SPS effect is restricted to the posterior electrodes. This strictly posterior distribution can be clearly seen in Fig. 4, showing the topographical distribution in the 500–750 ms and the 750–1000 ms latency windows, respectively. In addition, a negative shift preceding the P600/SPS is visible at frontal sites and Cz.

The overall ANOVAs did not yield a significant grammaticality effect for either of the two windows (500–750 ms: $F(1,31) < 1$; 750–1000 ms: $F(1,31)=1.53$, $MSe=13.21$, $P = 0.23$). However, the analyses per site showed a significant effect at PTL in the 500–750 ms window as well as significant effects at the PTL, Pz, and PTR sites in the 750–1000 ms window. The ANOVA testing for the topographical differences showed no significant window by electrode interaction ($F(1,31)=1.75$, $MSe=0.23$, $P > 0.05$). The ANOVA on the negativity in the 350–450 ms latency window showed no significant effects, neither overall nor in the analyses by electrode (All $F_s < 1$, other than Fz: ($F(1,31)=2.52$, $MSe=3.37$, $P = 0.12$) and ATR:

600/SPS, Subcategorization

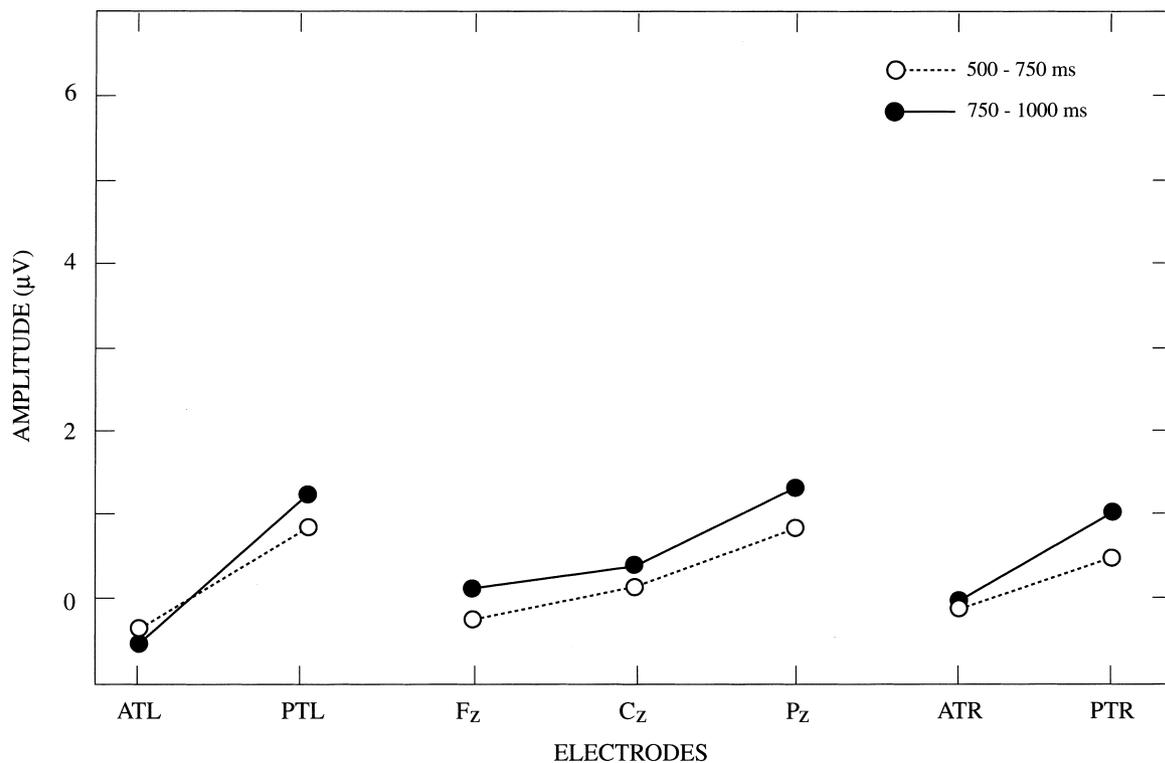


Fig. 4. Subcategorization condition, RSVP. The topographical distribution of the P600/SPS in the 500–750 ms latency window and in the 750–1000 ms latency window following presentation of the critical word (CW). For these two latency windows, the mean amplitude differences between the syntactically correct and the syntactically anomalous CW are displayed for two left hemisphere sites (ATL, PTL), three midline sites (Fz, Cz, Pz), and two right hemisphere sites (ATR, PTR).

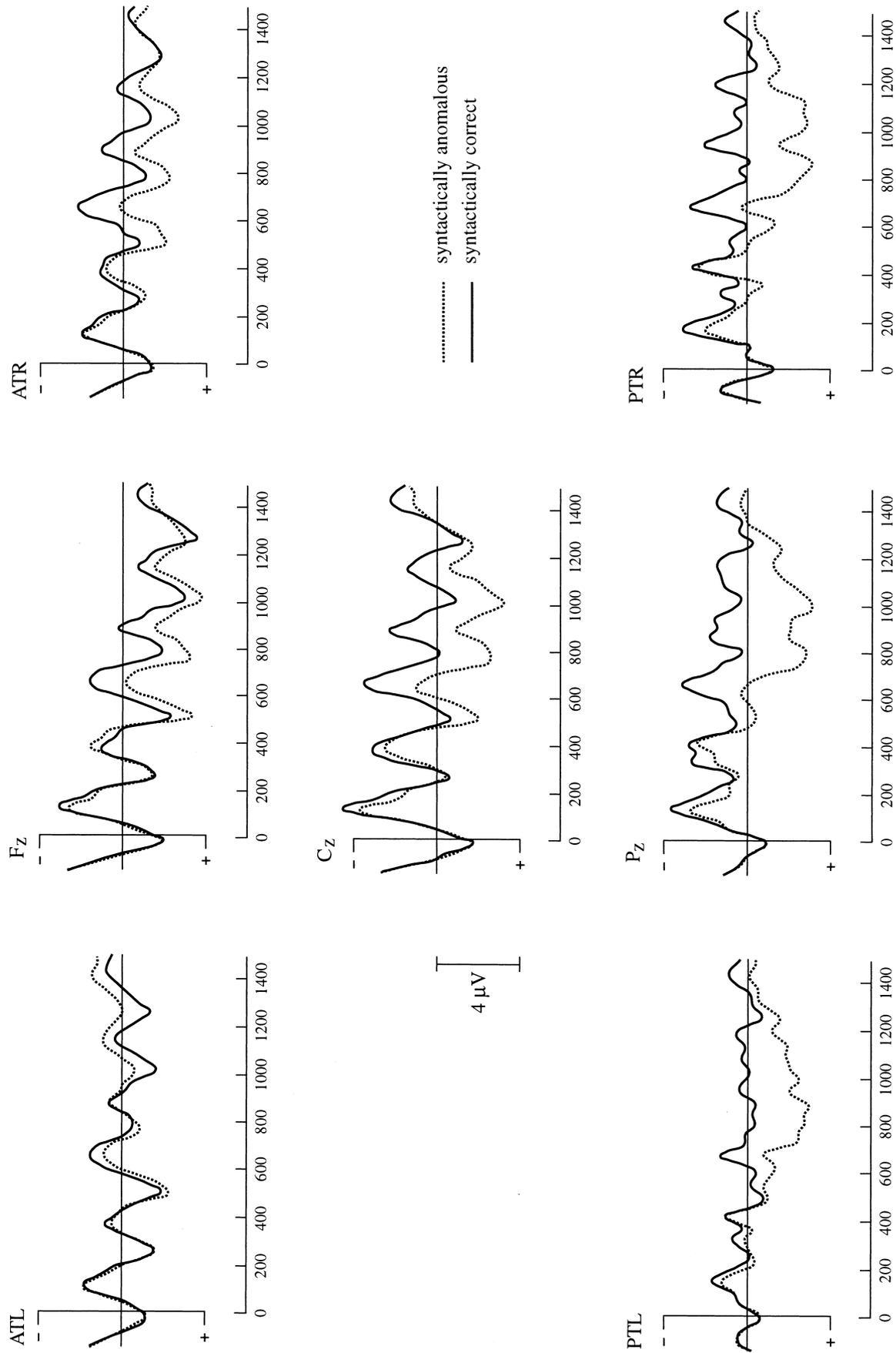


Fig. 5. Phrase structure condition, RSVP: Grand average ERPs (across all subjects and items), from 7 scalp sites, for the grammatically correct and incorrect critical words (CW). The onset of the CW is at zero milliseconds. The waveforms display the ERPs to the CW and the four following words. The time axis is in milliseconds. Each plot represents approximately 931 trials.

($F(1,31)=1.86$, $MSe=2.46$, $P = 0.18$). The presence of an early negativity preceding the P600/SPS could thus not be substantiated.

In short, the grammaticality effect of the subcategorization violation looks similar to the effect for the agreement violation in the 750–1000 ms window, but is clearly different in the 500–750 ms window.

3.2.3. Phrase structure

For the phrase structure violation a substantial positive shift is obtained (Fig. 5). This positive shift is visible over all electrode sites except ATL. Overall, a similar shift from a relatively equal anterior–posterior distribution in the 500–750 ms window to a more posterior distribution in the 750–1000 ms window is seen (Fig. 6). In addition the effect lasted longer over posterior than over anterior sites.

The ANOVAs for the two latency windows resulted in significant grammaticality effects (500–750 ms window: $F(1,31)=18.40$, $MSe=15.30$, $P < 0.0005$; 750–

1000 ms window: $F(1,31)=27.11$, $MSe=18.27$, $P < 0.0001$). However, only the analysis for the second window resulted in a significant interaction between grammaticality and site ($F(1,31)=14.60$, $MSe=1.79$, $P < 0.001$). This interaction was related to the increase of the effects over posterior sites in the second window (amplitude of the grammaticality effect averaged over the three frontal sites: 500–750 ms: $1.27 \mu\text{V}$; 750–1000 ms: $0.89 \mu\text{V}$; amplitude of the grammaticality effect averaged over the three posterior sites: 500–750 ms: $1.74 \mu\text{V}$; 750–1000 ms: $3.06 \mu\text{V}$). An ANOVA in which the topographical differences were directly tested after z-score normalization resulted in a significant window by electrode interaction ($F(1,31)=8.29$, $MSe=0.28$, $P < 0.01$). The statistical analyses thus support the conclusion that for the first window an equal distribution over the scalp was obtained, with a more posterior distribution for the second window. The ANOVA on the mean amplitudes in the 350–450 ms latency window did not result in significant effects, neither overall nor for individual electrode sites.

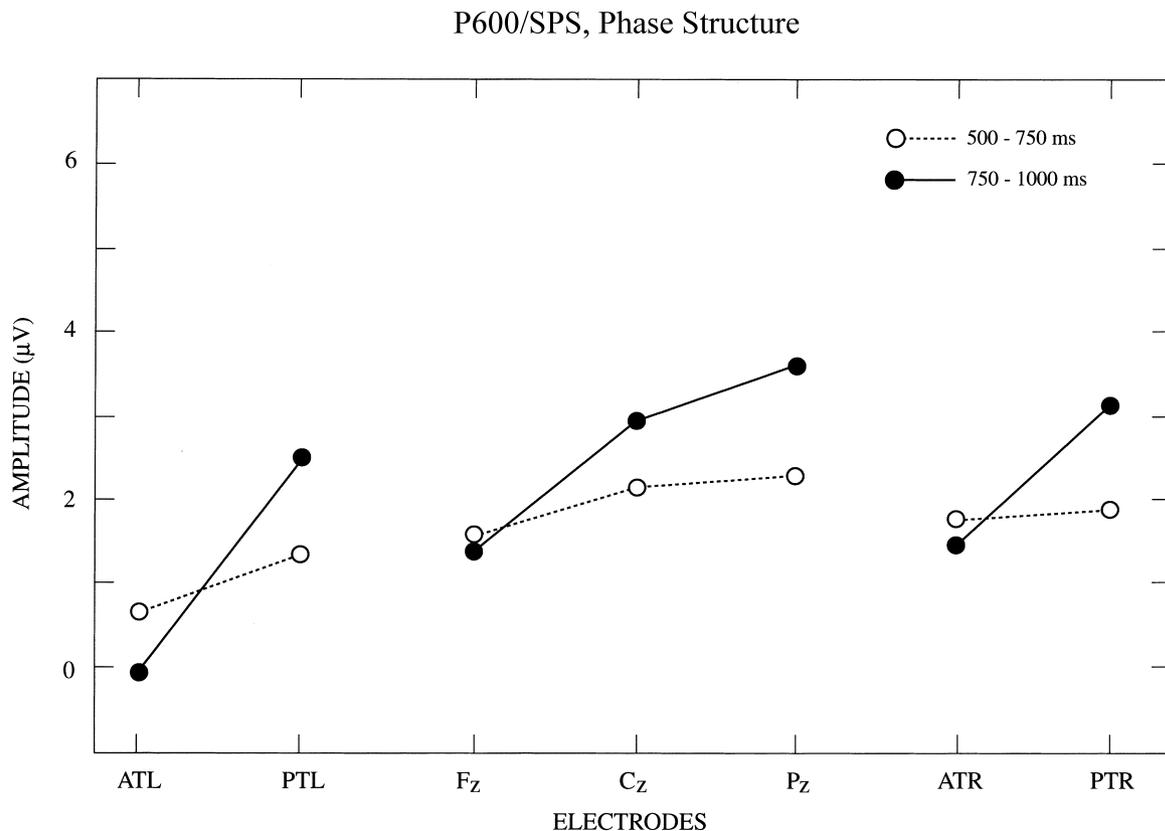


Fig. 6. Phrase structure condition, RSVP. The topographical distribution of the P600/SPS in the 500–750 ms latency window and in the 750–1000 ms latency window following presentation of the critical word (CW). For these two latency windows, the mean amplitude differences between the syntactically correct and the syntactically anomalous CW are displayed for two left hemisphere sites (ATL, PTL), three midline sites (Fz, Cz, Pz), and two right hemisphere sites (ATR, PTR).

3.3. Discussion

Four aspects of the results are noteworthy. First, even with a presentation rate of four words per second, violations of syntactic constraints and of preferred structural assignments elicit substantial positive shifts in the ERP waveforms. Interestingly, the onset of the positive shift is not different from what we observed when a much slower rate of presentation was used [11]. This result, then, substantiates the relevance of our earlier findings for characterizing the on-line processing of syntactic information during reading.

A second aspect worth mentioning is that the data seem to indicate that the P600/SPS actually consists of two parts, one part with a fairly equal distribution along the anterior–posterior axis, and a second part with a clear posterior distribution and a longer duration over posterior sites.

Third, in contrast to our earlier studies [10,11] a positive shift was observed for the subcategorization violation. However, this effect seemed to consist largely of the second part of the P600/SPS. The findings for the subcategorization violation suggest that the two aspects of the P600/SPS can be dissociated. In this particular case, it might be that the inherent linkage between semantic and syntactic features embodied in the argument structure of verbs (cf. [3,10]) resulted in a sentence interpretation problem for both semantic and syntactic reasons. Therefore the overlap of syntactic effects and their semantic consequences might have affected the overall shape of the effect, for example, by a partial overlap of N400 and P600/SPS effects.

Finally, in contrast to our earlier study with a slower presentation rate, the P600/SPS was not immediately followed by an increased N400 to words that followed the violation. With the rapid presentation rate the semantic integration processes might be shifted in relation to the rate at which new information comes in. We will postpone further discussion about the functional interpretation of these results until after presenting the second experiment with speech as input.

4. Experiment 2: The P600/SPS in spoken sentences

4.1. Method

The exact same sentences were used as in the previous experiment. All materials were spoken by a female speaker and further processed in the same way as in Experiment 1. The resulting two experimental lists together with practice and filler sentences were recorded onto DAT-tape.

4.1.1. Grammaticality judgement pre-test

We first tested the sentences in a grammaticality jud-

gement experiment, using a Go/NoGo task, in which subjects were instructed to respond whenever they detected a grammatical violation. The purpose of this pre-test was to evaluate whether the violations are as easily detected in spoken sentences as they were in their written form [11].

The two test tapes were played to 15 subjects each. The subjects were instructed to listen to each sentence for comprehension, and to press a button whenever they encountered a grammatical error. The correctness of the response was emphasized rather than the speed. No information was given concerning the kinds of errors that would occur. The onsets of the critical words were marked by a trigger pulse that started a timer, that was stopped by the pushbutton response. The response times were stored by computer for off-line analysis.

The results for sentences with a grammatical violation are presented in Table 1, for each violation type separately. Table 1 specifies both the percentage of correct rejections for the ungrammatical sentences and the mean response times. False alarms on correct sentences occurred on at most 3% of the grammatically correct items for a particular violation type. For comparison, Table 1 also presents the percentages of correct responses obtained in the visually presented grammaticality pretest [11].

As is clear from Table 1, subjects pick up on the ungrammaticalities in the spoken sentences as easily as in the written version of the same sentences. The response latencies are fairly similar for the three violation types and occur within almost one second after the onset of the critical word. We take this as evidence for the validity of the critical word as the point in the sentence where the assignment of the preferred syntactic structure becomes impossible.

Table 1

Percentage of responses (and mean response times) by violation type, relative to the onset of the critical word (CW), for both visual and auditory presentation

	Visual presentation (derived from Ref. [11])	Auditory presentation
Agreement	97%	97% (1034 ms)
Subcategorization	91%	97% (1011 ms)
Phrase Structure	99% ^a	98% (1131 ms)

^a The position of the CW in the Hagoort et al. (1993) study [11] was the noun in the NP. In the present experiment, on the basis of the results of the earlier studies, we defined the adverb preceding the noun in the ungrammatical sentences as the Critical Word. Therefore the pre-CW responses in Table 1 of the Hagoort et al. study are added to the responses on the CW and following words.

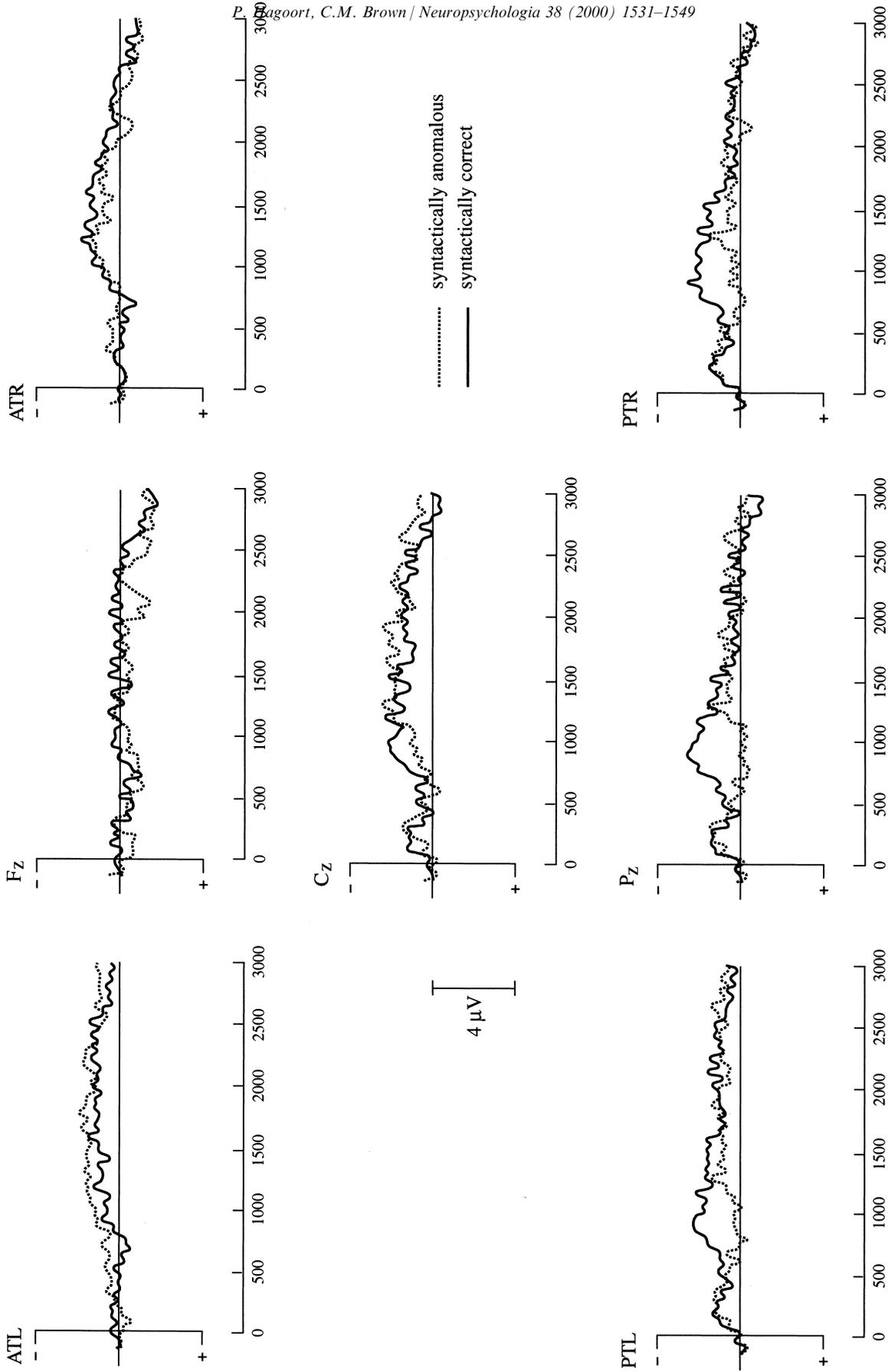


Fig. 7. Agreement condition, connected speech: Grand average ERPs (across all subjects and items), from 7 scalp sites, for the grammatically correct and incorrect critical words (CW). The onset of the CW is at zero milliseconds. The time axis is in milliseconds. Each plot represents approximately 786 trials.

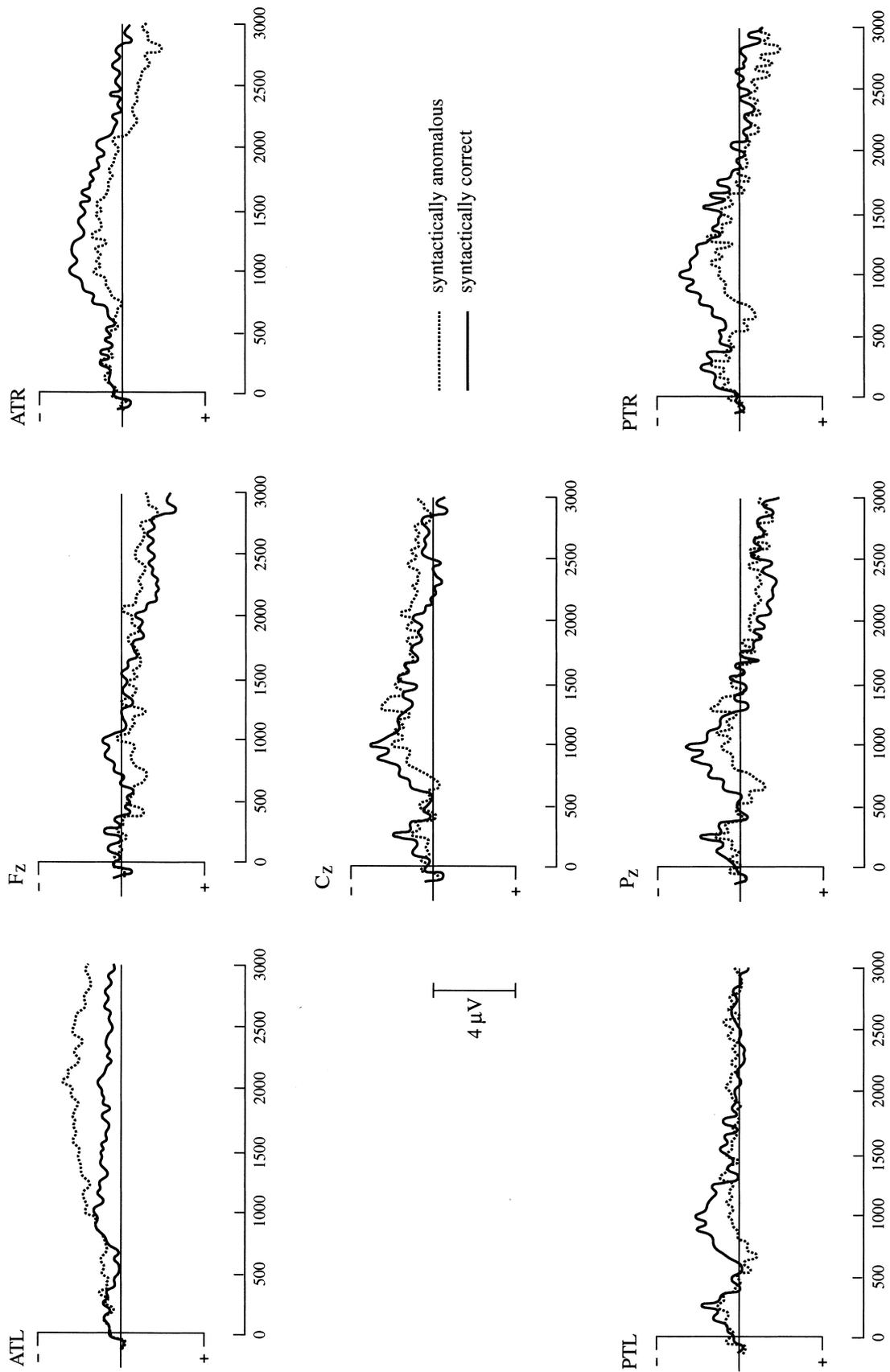


Fig. 8. Subcategorization condition, connected speech: Grand average ERPs (across all subjects and items), from 7 scalp sites, for the grammatically correct and incorrect critical words (CW). The onset of the CW is at zero milliseconds. The time axis is in milliseconds. Each plot represents approximately 786 trials.

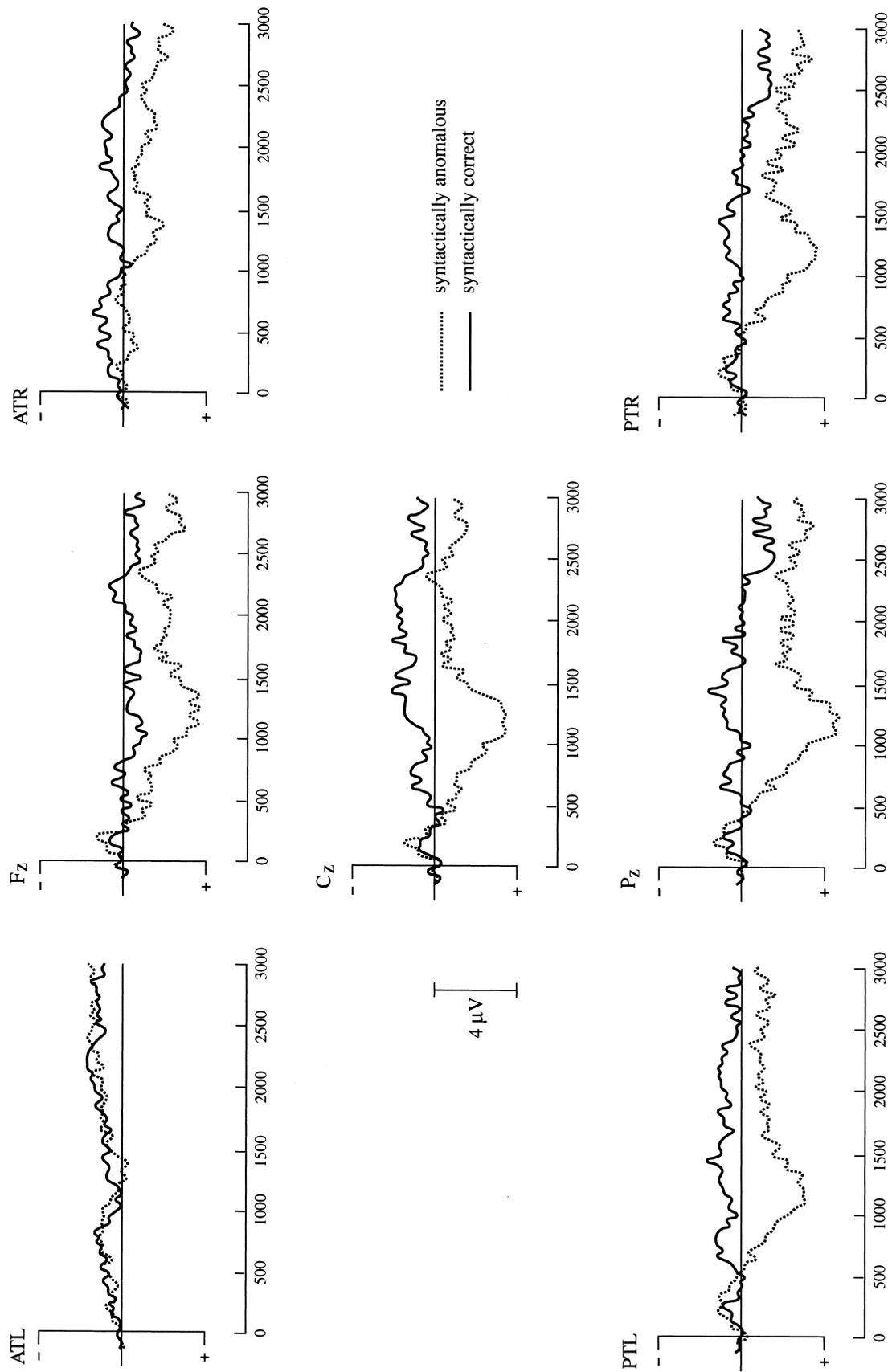


Fig. 9. Phrase structure condition, connected speech: Grand average ERPs (across all subjects and items), from 7 scalp sites, for the grammatically correct and incorrect critical words (CW). The onset of the CW is at zero milliseconds. Each plot represents approximately 786 trials.

4.1.2. Subjects

A total of 30 subjects (23 female; 30 right-handers) participated in the ERP experiment, 15 on each list.

4.1.3. Procedure

Subjects were asked to fixate on the center of a display unit on the monitor in front of them, while the spoken sentences were presented over headphones. The average length of the sentences was 4.17 s, with a shortest sentence of 2.99 sec and a longest sentence of 6.06 s. After sentence offset there was a variable delay of at least 500 ms (for the longest sentence), followed by the presentation of an asterisk for 2 s. Subjects were told to blink during this period. The presentation of the asterisk was followed by another 500 ms pause before the next sentence started.

The subjects were instructed to listen to the sentences for understanding. No additional task demands were imposed. The subjects were told that some sentences would be grammatically incorrect, but they were given no information concerning the kinds of grammatical errors that would occur. The entire session, including electrode application and removal lasted approximately 1.5 h.

4.2. Results

After removal of the trials with artifacts (12.7%), average waveforms were computed in exactly the same way as in the previous experiment. Figs. 7, 8 and 9 show the waveforms for the three violation types. All three figures show the waveforms from the acoustic onset of the critical word, for a period of 3 s, which for most items includes the rest of the sentence. The overall shape of the waveforms in these figures is clearly different from the overall morphology of the waveforms in the preceding experiment. This difference is due to the characteristics of connected speech, which, in contrast to the same sentences in their orthographic presentation, lacks clear physical boundaries between the words in the sentences.

Most importantly, however, just as for their written counterparts the spoken sentences elicited a clear P600/SPS for all three violation types. This effect starts at about 500 ms after the onset of the critical word and lasts for a substantial period of time, until about 1250 ms after CW-onset and even longer for the phrase structure violation. The effect shows a clear posterior distribution for all three violation types, although especially for the phrase structure condition the distribution between the early and the later windows seems to be slightly different. Over posterior sites the effect is roughly symmetric. However, anterior sites show some asymmetry in the effects of grammaticality.

Repeated measures ANOVAs were performed for each of the three violation types separately. Analyses

were based on the mean amplitudes relative to a 150 ms baseline preceding the critical word in three latency windows of 250 ms each: 500–750 ms, 750–1000 ms, and 1000–1250 ms after onset of the critical word. The size of the latency windows thus matched that of the windows in the RSVP experiment. The relevant factors in the ANOVAs were grammaticality, window, and site. The Greenhouse–Geisser correction was applied as before. Since main effects of window and electrode are without any theoretical importance for this study, they will not be reported. Differences in topographical distributions of the three latency windows were tested in the same way as in the first experiment.

4.2.1. Agreement

Fig. 7 shows an effect of grammaticality that is mainly restricted to the posterior electrode sites. The positive shift for the ungrammatical condition starts at 500 ms after the acoustic onset of the word that renders the sentence ungrammatical. It becomes maximal between 750 and 1000 ms, and lasts until about 1250 ms post-CW. In addition to the positive shift, the frontal left and right electrode site showed an earlier negative shift between about 300 and 550 ms. This frontal negativity was also reported by Osterhout and Holcomb [24]. However, the latency of this effect was clearly later than the ELAN reported by Hahne and Friederici [13].

An ANOVA with all electrode sites failed to show a significant effect of grammaticality ($F(1,29)=2.09$, $MSe=65.32$, $P > 0.10$). However, the grammaticality by site interaction was significant ($F(1,29)=6.14$, $MSe=5.23$, $P < 0.05$), due to the posterior topography of the effect. We therefore performed an additional ANOVA with the inclusion of only the three posterior sites (PTL, Pz, PTR). This analysis resulted in a significant effect (1.44 μV) of grammaticality ($F(1,29)=6.90$, $MSe=40.59$, $P < 0.05$). The largest effect was observed in the 500–750 ms window (2.05 μV). After z -score normalization, the window by electrode interaction failed to reach significance ($F < 1$), substantiating the absence of topographic differences in the three latency windows. The statistical results for the early negativities are summarized in a separate section below.

4.2.2. Subcategorization

The subcategorization violation also elicited an effect of grammaticality, as can be seen in Fig. 8. The effect is largest over the posterior sites. A clear asymmetry is visible over more frontal sites, where ATL shows a negativity between about 400 and 700 ms to the violation, and a late negativity from 900 ms onwards, whereas ATR shows a long lasting positive shift, start-

ing at 500 ms. As for the agreement violation, the positive shift is maximal in the 750–1000 ms window.

An overall ANOVA just failed to show a significant effect of grammaticality ($F(1,29)=3.64$, $MSe=54.56$, $P=0.07$). The interaction between grammaticality and site also failed to reach significance ($F(1,29)=3.89$, $MSe=4.22$, $P>0.05$). The ANOVA on the posterior electrode sites (PTL, Pz, PTR), however, resulted in a significant effect (1.22 μV) of grammaticality ($F(1,29)=8.14$, $MSe=24.80$, $P<0.01$), with the largest effect (1.57 μV) in the 750–1000 ms window. After z -score normalization, the window by electrode interaction failed to reach significance ($F(1,29)=2.09$, $MSe=0.26$, $P>0.05$), substantiating the absence of topographic differences in the three latency windows.

4.2.3. Phrase structure

As in the previous experiment, the largest grammaticality effect is obtained for the phrase structure violation (see Fig. 9). Again the effects are strongest over posterior sites. As for the written sentences, no positive shift is seen over the left frontal site (ATL), but over the right frontal site (ATR) such a shift is clearly present. At Fz and ATR the effect seemed to be somewhat earlier than over centroparietal sites. In contrast to the other violation types the effect is largest in the 1000–1250 ms window. This is most likely due to the fact that the sentences in the phrase structure condition became ungrammatical at the word following the critical word, which thereby might give an extra boost to the positive shift elicited to the critical word (i.e. the adverb preceding the noun). As was the case for the P600/SPS in the RSVP experiment, apart from ATL the initial part of the P600/SPS (500–750 ms) shows a more equal anterior–posterior distribution than the later part, which is clearly maximal over posterior sites.

The overall ANOVA yielded a significant effect of grammaticality ($F(1,29)=15.18$, $MSe=73.36$, $P<0.001$), as well as a significant interaction between grammaticality and site ($F(1,29)=8.56$, $MSe=6.32$, $P<0.01$), and a significant interaction between grammaticality and window ($F(1,29)=12.02$, $MSe=6.37$, $P<0.01$). The ANOVA on the three posterior electrode sites likewise resulted in a highly significant effect (2.57 μV) of grammaticality ($F(1,29)=19.00$, $MSe=47.00$, $P=0.0001$), and a significant grammaticality by window interaction ($F(1,29)=33.43$, $MSe=2.74$, $P<0.0001$), with the largest effect (4.06 μV) in the 1000–1250 ms window. After z -score normalization, the ANOVA resulted in a significant window by electrode interaction ($F(1,29)=7.35$, $MSe=0.23$, $P<0.02$). As in the RSVP version, the distribution became more posterior in the second part of the effect.

4.2.4. Left frontal negativities

In two of the violation types, the agreement and the subcategorization violations, frontal negativities were seen that preceded the P600/SPS in latency. For the agreement violation this early negativity was visible at both the ATL and the ATR site. For the subcategorization violation the negativity emerged only at ATL. A left frontal negativity in relation to syntactic violations has been reported before [7], most prominent with speech as input [24]. Although ATL is not as frontal as F7, the left frontal negativity has also been observed over this electrode site [7]. In our experiment the onset of this frontal negativity was at about 350 ms. We tested the left frontal negativity (LAN) for these two violation types in an ANOVA on the mean amplitudes for the ATL site in a latency window of 350–550 ms after onset of the critical word (cf. [7]). The relevant factors in the ANOVA were violation type (agreement, subcategorization), and grammaticality (grammatical, ungrammatical). This analysis resulted in a marginally significant effect of grammaticality ($F(1,29)=3.74$, $MSe=1.64$, $P=0.06$). The interaction between violation type and grammaticality was not significant ($F<1$). The phrase structure violation did not result in a left frontal negativity.

4.3. Discussion

The results of this experiment show that listening to sentences spoken at a normal rate and with normal intonation elicits a substantial P600/SPS time locked to the word that renders the preferred structural assignment impossible. The electrophysiological response to visual and auditory syntactic violations is very similar, but modality-related differences were also obtained. The similarity across modalities resides in the polarity and the latency of the effects. It is remarkable that across presentation rates and modalities the onset of the P600/SPS remained almost identical. Fig. 10 illustrates this for the P600/SPS elicited by the same agreement violation at a presentation rate of 1.6 words per second, 4 words per second, and with speech. As can be seen in this figure, in all three cases the onset of the effect was in the order of 500 ms after onset of the critical word in the sentence. This across-modality consistency in the onset latency of the P600/SPS is in agreement with the results of Friederici et al. [7], but stands in clear contrast to the findings of the Osterhout and Holcomb study [24]. These authors reported earlier P600/SPS effects for spoken input compared to visual input. It is currently unclear what causes this discrepancy between their and our studies or what might modulate the onset of the P600/SPS.

The results of our study are, however, clearly compatible with the outcome of the Osterhout and Holcomb study with respect to the topography of the

P600/SPS in the auditory modality. As was observed by these authors, we also found a more posterior distribution for the auditory P600/SPS compared to the visual P600/SPS. In addition, we replicated the finding of (left) frontal negativities for at least two of the syntactic violations in our experiment. The studies with English [24], German [7] and Dutch (this study) together, then, present clear evidence that different syntactic violations in different languages result in a very similar pattern of effects. The finding that the P600/SPS occurs not only in reading but also while listening to speech supports the claim that core aspects of parsing operations are identical across the two modalities of input.

The results for the phrase structure violation suggest that the auditory P600/SPS might also consist of an earlier and a later phase with similar distributional characteristics as were observed for fast reading. However, given the durational differences between the critical words in the auditory modality, the overall ERP waveforms most likely show more overlap of these two parts of the P600/SPS complex in the auditory modality than in the visual modality.

5. General discussion

For two of the three syntactic violations in Experiment 2 we observed a relatively early negativity, with a (left) frontal maximum. This anterior negativity was also observed by Osterhout and Holcomb [24] in their study on syntactic violations in speech. Recently, Friederici et al. [7] have claimed that the early left frontal

negativities are elicited by word category violations. That is, if on the basis of the syntactic context a particular word class (e.g. noun, verb, adjective, etc.) is allowed whereas other word classes are not, the early negativity is seen to a word of a category that is illegal in that position. For instance, if on this account a noun occurs at a position in the sentence where a verb is required, not only a P600/SPS but also the early negativity should be observed. The underlying idea is that word category information is crucial for the on-line first-pass assignment of syntactic structure. In the view of Friederici et al. [7], the early left frontal negativity is therefore a sensitive measure for first-pass parsing operations (for further details and discussion, see [7]).

This interesting proposal is, however, not corroborated by the data of this study. For two syntactic violations an early frontal negativity was observed. However, only one of these embodied a word category violation. In the subcategorization violation an article was presented where a preposition was expected or required (e.g. *‘The boy boasts the car of his father.’). The agreement violation, in contrast, contained the correct word category (i.e. a verb) but with the wrong marker for number (e.g. *‘The boy throw the toy on the ground.’). This is a morphosyntactic error, but clearly not a word category violation. Nevertheless, a left frontal negativity was observed to the verb with the incorrect number marking. Finally, in the phrase structure violation, the adverb following the adjective resulted in an P600/SPS, but not in a left frontal negativity. This violation, however, can be construed as a violation of the preferred word category, since most

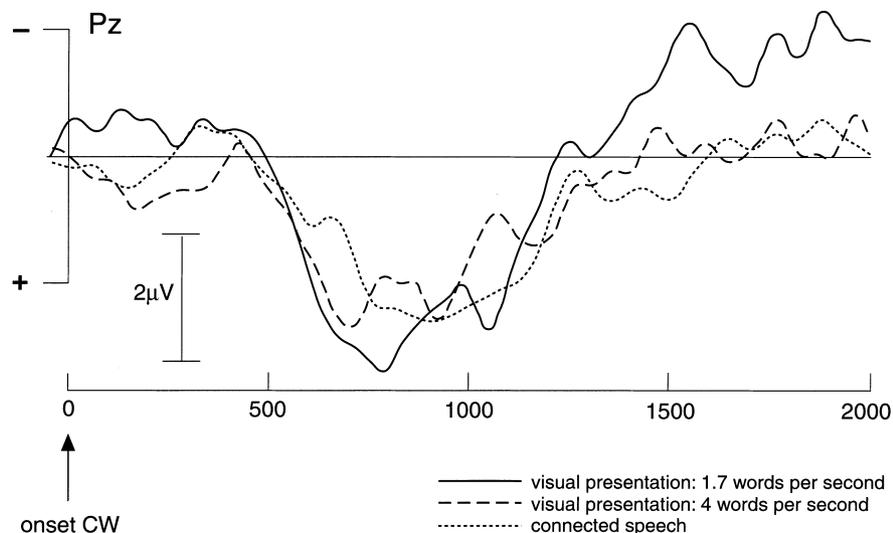


Fig. 10. The P600/SPS with speech, slow and fast visual presentations, for the agreement condition. The difference ERP waveforms at electrode Pz are displayed for connected speech, a visual presentation of 1.7 words per second, and a visual presentation of 4 words per second. The onset of the critical word (CW) is at zero milliseconds. The time axis is in milliseconds.

likely a noun was expected at the position where the adverb occurred (cf. [11]). Together, this pattern of results is at odds with the proposal of Friederici et al. [7].

To date a number of studies have reported (left) anterior negativities in relation to syntactic processing [6,7,16,20,21,30]. However, at present the variation in latency and exact topography, as well as in the conditions that elicit this ERP effect prevent the conclusion that all these findings can be subsumed under the heading of one ERP component or even one ERP effect (cf. [2,12]). More fine-grained experimentation is required to determine the functional significance of these frontal negativities.

The main outcome of this study is that we obtained P600/SPS effects to syntactic violations in written and spoken sentences. The similarity of the effects for reading and for listening supports the claim that reading and listening share central aspects of postlexical sentence processing.

In contrast to other reports in the literature [24], the onset latency of the P600/SPS was very similar in both modalities. However, this similarity does not necessarily imply that the underlying processes have the same time course in both modalities. The crucial difference between reading and the processing of speech is the difference in the time at which word information is made available. In reading, words are essentially instantaneously available, whereas in speech the information accrues in a left-to-right order. Therefore, relative to the word recognition point of the critical words in our experiments, the onsets of the auditory N400 and the auditory P600/SPS are actually relatively early.

This study also provided evidence that the P600/SPS might in fact reflect two aspects of the parsing process, which especially in Experiment 1 showed up as an early phase between about 500 and 750 ms with a fairly similar distribution along the anterior–posterior axis, and a later phase after 750 ms with a clear posterior distribution and a longer duration over posterior sites. Although these two aspects of the P600/SPS have not been reported before, reanalysis of our earlier data [11] and inspection of the waveforms of other P600/SPS studies [26,28], clearly support our observation in this study of the P600/SPS being a complex consisting of more than one aspect of the parsing process. We hypothesize that one aspect is related to the complexity of syntactic processing. Whenever the complexity of the structural integration process increases, a P600/SPS is observed [15]. This account is supported by the finding that whenever the syntactic manipulation is not in the form of a straightforward violation but implicates differences in syntactic complexity [1,12] the distribution of the P600/SPS is more frontal than in the case of a straightforward grammatical violation. In the

latter case a more posteriorly oriented distribution of the P600/SPS is observed which might indicate the failure of a parse, and/or the attempt at a revision of the syntactic structure. We suggest that in the current experiments the first phase related to increased syntactic integration difficulty, which was followed by a second phase indicating the failure/revision of the parse [5]. Clearly more research is needed to test these ideas.

A clear separation between the distribution of the two aspects of the P600/SPS was not observed for all conditions in this study. In the RSVP experiment the subcategorization condition did not show the first phase as strongly as the other violations. On the basis of previous findings [10,11] we have suggested that this violation is both syntactic and semantic in nature, with a resulting overlap of the P600/SPS and an N400 effect. It could be that in the timing conditions of the RSVP experiment the N400 effect had its overlap most strongly with the first phase of the P600/SPS. As a result only the more posterior phase remained clearly visible. In the spoken sentences, overlap between first and second phase is to be expected since the spoken syntactic violations are stretched out over time and vary in their duration. As a result it becomes more difficult to obtain clearly non-overlapping phases in the averaged waveforms. Nevertheless, separable phases were found for the phrase structure violation.

References

- [1] van Berkum JJA, Brown CM, Hagoort P. Early contextual effects in sentence processing: evidence from event-related brain potentials. *Journal of Memory and Language* 1999;41:147–82.
- [2] Brown CM, Hagoort P, Kutas M. Postlexical integration processes in language comprehension: evidence from brain-imaging research. In: Gazzaniga MS, editor. *The new cognitive neurosciences*. 2nd ed. Cambridge MA: The MIT Press, 2000. p. 881–95.
- [3] Fisher C, Gleitman H, Gleitman LR. On the semantic content of subcategorization frames. *Cognitive Psychology* 1991;23:331–92.
- [4] Fodor JA. *The modularity of mind*. Cambridge MA: The MIT Press, 1983.
- [5] Fodor JD, Inoue A. The diagnosis and cure of garden paths. *Journal of Psycholinguistic Research* 1994;5:407–34.
- [6] Friederici AD, Pfeifer E, Hahne A. Event-related brain potentials during natural speech processing: effects of semantic, morphological, and syntactic violations. *Cognitive Brain Research* 1993;1:183–92.
- [7] Friederici AD, Hahne A, Mecklinger A. Temporal structure of syntactic parsing: early and late event-related brain potential effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 1996;22:1219–48.
- [8] Ganis G, Kutas M, Sereno MI. The search for ‘common sense’: an electrophysiological study of the comprehension of words and pictures in reading. *Journal of Cognitive Neuroscience* 1996;8:89–106.
- [9] Greenhouse SW, Geisser S. On methods in the analysis of profile data. *Psychometrika* 1959;24:95–112.
- [10] Hagoort P, Brown C. Brain responses to lexical ambiguity resol-

- ution and parsing. In: Clifton Jr C, Frazier L, Rayner K, editors. *Perspectives on sentence processing*. Hillsdale NJ: Erlbaum, 1994. p. 45–80.
- [11] Hagoort P, Brown C, Groothusen J. The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes* 1993;8:439–83.
- [12] Hagoort P, Brown CM, Osterhout L. The neurocognition of syntactic processing. In: Brown CM, Hagoort P, editors. *The neurocognition of language*. Oxford UK: Oxford University Press, 1999. p. 273–316.
- [13] Hahne A, Friederici AD. Electrophysiological evidence for two steps in syntactic analysis: early automatic and late controlled processes. *Journal of Cognitive Neuroscience* 1999;11:194–205.
- [14] Jasper HH. Report to the committee on methods of clinical examination in electroencephalography. Appendix: the 10–20 system of the International Federation. *Electroencephalography and Clinical Neurophysiology* 1958;10:371–5.
- [15] Kaan E, Harris A, Gibson E, Holcomb P. The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes* 2000;15:159–201.
- [16] Kluender R, Kutas M. Subjacency as a processing phenomenon. *Language and Cognitive Processes* 1993;8:573–633.
- [17] Kutas M, Hillyard SA. Reading senseless sentences: brain potentials reflect semantic anomaly. *Science* 1980;207:203–5.
- [18] Kutas M, Van Petten CK. Psycholinguistics electrified: event-related brain potential investigations. In: Gernsbacher MA, editor. *Handbook of psycholinguistics*. San Diego: Academic Press, 1994. p. 83–143.
- [19] McCarthy G, Wood CC. Scalp distributions of event-related potentials: An ambiguity associated with analysis of variance models. *Electroencephalography and Clinical Neurophysiology* 1985;62:203–8.
- [20] Münte TF, Heinze HJ, Mangun GR. Dissociation of brain activity related to syntactic and semantic aspects of language. *Journal of Cognitive Neuroscience* 1993;5:335–44.
- [21] Neville H, Nicol JL, Barss A, Forster KI, Garrett MF. Syntactically based sentence processing classes: evidence from event-related brain potentials. *Journal of Cognitive Neuroscience* 1991;3:151–65.
- [22] Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 1971;9:97–113.
- [23] Osterhout L, Holcomb PJ. Event-related potentials elicited by syntactic anomaly. *Journal of Memory and Language* 1992;31:785–806.
- [24] Osterhout L, Holcomb PJ. Event-related potentials and syntactic anomaly: Evidence of anomaly detection during the perception of continuous speech. *Language and Cognitive Processes* 1993;8:413–37.
- [25] Osterhout L, Holcomb PJ. Event-related potentials and language comprehension. In: Rugg MD, Coles MGH, editors. *Electrophysiology of mind: event-related brain potentials and cognition*. Oxford, UK: Oxford University Press, 1995.
- [26] Osterhout L, Mobley LA. Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language* 1995;34:739–73.
- [27] Osterhout L, Holcomb PJ, Swinney DA. Brain potentials elicited by garden-path sentences: evidence of the application of verb information during parsing. *Journal of Experimental Psychology: Learning, Memory and Cognition* 1994;4:786–803.
- [28] Osterhout L, McKinnon R, Bersick M, Corey V. On the language-specificity of the brain response to syntactic anomalies: is the Syntactic Positive Shift a member of the P300 family? *Journal of Cognitive Neuroscience* 1996;8:507–26.
- [29] Patel AD, Gibson E, Ratner J, Besson M, Holcomb PJ. Processing syntactic relations in language and music: an event-related potential study. *Journal of Cognitive Neuroscience* 1998;10:717–33.
- [30] Rösler F, Friederici AD, Pütz P, Hahne A. Event-related brain potentials while encountering semantic and syntactic constraint violations. *Journal of Cognitive Neuroscience* 1993;5:345–62.
- [31] Rösler F, Heil M, Glowalla U. Monitoring retrieval from long-term memory by slow event-related brain potentials. *Psychophysiology* 1993;30:170–82.
- [32] Winer BJ. *Statistical principles in experimental design*. 2nd ed. New York: McGraw-Hill, 1971.