ACADEMIC
PRESS

Available online at www.sciencedirect.com
SCIENCE DIRECT®

Brain and Language 86 (2003) 142–164

Brain
and
Language

www.elsevier.com/locate/b&l

Electrophysiological correlates of prosody and punctuation☆

Karsten Steinhauer*

Department of Neuroscience, Georgetown University, Washington, DC, USA

Accepted 23 July 2002

Abstract

Psycholinguistic models of sentence parsing are primarily based on reading rather than auditory processing data. Moreover, both prosodic information and its potential orthographic equivalent, i.e., punctuation, have been largely ignored until recently. The unavailability of experimental online methods is one likely reason for this neglect. Here I give an overview of six event-related brain potential (ERP) studies demonstrating that the processing of both prosodic boundaries in natural speech and commas during silent reading can determine syntax parsing immediately. In ERPs, speech boundaries and commas reliably elicit a similar online brain response, termed the *Closure Positive Shift* (CPS). This finding points to a common mechanism, suggesting that commas serve as visual triggers for covert phonological phrasing. Alternative CPS accounts are tested and the relationship between the CPS and other ERP components, including the P600/SPS, is addressed.

Keywords: Event-related potentials; Language; Prosody; Punctuation; Syntax; Commas; P600; Closure positive shift; CPS; Garden path sentences

1. Introduction

According to Chafe (1988) many people are likely to hear some kind of 'inner voice' when silently reading texts such as a letter—or this paper. This phenomenon reflects the introspective experience of written words activating their corresponding phonological representations and is usually referred to as phonological re-coding (e.g., Share, 1999). As a consequence, even reading studies on lexical processing

---

[☆] The six experiments described in this paper were conducted at the Max Planck Institute of Cognitive Neuroscience. The work was supported by a grant from the German Research Foundation (DFG, FR519/17) and by the Leibniz Science Prize awarded to Angela D. Friederici by the German Research Foundation. I thank Kai Alter for recording and processing the speech signals, Martin Meyer, Lyn Frazier, Janet Fodor, and Atsu Inoue for helpful discussions, Cornelia Schmidt, Ina Koch, and Andrea Gast-Sandmann for their skilful technical support. I am also grateful to Laurie Stowe, Clare Foa, Galen Clayton, and an anonymous reviewer for important comments on an earlier version of the manuscript.

[*] Fax: 1-202-687-6914.
E-mail address: ks42@georgetown.edu.

doi:10.1016/S0093-934X(02)00542-4

normally control for phonological similarities and neighborhood effects among the lexical entries. However, the inner voice appears to provide richer information than the pure sound pattern of words. It also seems to comprise supra-segmental phonology including sentence accents and intonational phrasing, which is not unlikely to correspond to certain prosodic patterns in spoken language. This 'covert,' 'subvocal,' or 'implicit' prosody during reading appears to be influenced by punctuation. A number of previous studies have suggested that overt prosody in spoken language (Marslen-Wilson, Tyler, Warren, Genier, & Lee, 1992; Speer, Kjelgaard, & Dobbroth, 1996; Warren, Grabe, & Nolan, 1995) and both punctuation (Clifton, 1993; Cohen, Douaire, & Elsabbagh, 2001; Hill & Murray, 2000 and unpublished data; Mitchell & Holmes, 1985) and implicit prosody during reading (Bader, 1998; Fodor, 1998) may have a strong influence on how we comprehend sentences. In particular, this non-lexical information may guide syntactic parsing and thereby prevent, or cause, initial misunderstandings, i.e., garden path phenomena. Unfortunately, in speech research data collection with behavioral online methods seemed to be prone to strategic influences, and many of these studies have been subject to a variety of criticisms due to methodological weaknesses (Cutler, Dahan, & van Donselaar, 1997; Warren, 1999; Watt & Murray, 1996). As a consequence, most current models of sentence parsing, including our understanding of parsing preferences and garden path effects, rest predominantly on much easier-to-control reading studies that do not require time consuming acoustic analyses. In this domain, punctuation has usually been either omitted entirely or restricted to so-called 'unambiguous' baseline conditions, thus simply implying the disambiguating potential of punctuation rather than investigating its mechanisms.

In a recent paper, Warren (1999) suggested that new methods such as event-related potentials (ERPs) might be needed to shed new light on the mental operations involved in prosodic processing. ERPs reflect the real-time electrophysiological brain activity of cognitive processes that are time-locked to the presentation of target stimuli. The method is non-invasive and data collection does not require complex task performance or the unnatural interruption of stimulus presentation. In psycholinguistic research, distinct ERP components for semantic and syntactic processes have been identified. Difficulties in lexical processing and conceptual–semantic integration elicit the N400 component, a centro-parietal negativity peaking between 400 and 600 ms after onset of the target word (Chwilla, Brown, & Hagoort, 1995; Kutas & Hillyard, 1980). Difficulties in syntactic processing due to ungrammatical violations yield early (150–500 ms) left anterior negativities (LANs) which have been linked to automatic computations (Friederici, Steinhauer, & Frisch, 1999; Gunter, Stowe, & Mulder, 1997; Hahne & Friederici, 1999). Syntactic processing difficulties due to violations as well as garden path effects and complex structures also elicit late (600–1000 ms) centro-parietal positivities, referred to as P600 components or syntactic positive shifts (SPS) (Friederici, Steinhauer, Mecklinger, & Meyer, 1998; Hagoort, Brown, & Groothusen, 1993; Osterhout & Holcomb, 1992). Unlike the LAN, the P600 seems to be associated with controlled processes of structural reanalysis and repair (Hahne & Friederici, 1999) and may comprise distinct subcomponents (Friederici, Mecklinger, Spencer, Steinhauer, & Donchin, 2001). Whereas LAN, N400, and P600 components for syntactic and semantic processing were replicated across different languages, including an artificial language (Friederici, Steinhauer, & Pfeifer, 2002), ERP components reflecting phonological processing have been observed only occasionally. Connolly and Phillips (1994) reported a phonological mismatch negativity (PMN) for sentence terminal words that differed phonologically from the expected target word. After insertion of an artificial pause into spoken sentences, Besson, Faita, Czternasty, and Kutas (1997) found both 'emitted potentials'[1] as well as enhanced amplitudes for the so-called N100–P200 complex of the delayed next word. In a Dutch study on metrical stress patterns, Böcker, Bastiaansen, Vroomen, Brunia, and Gelder (1999) observed that bi-syllabic words with irregular stress on the second syllable elicited an N325 component. Although these studies demonstrate that basic phonological processing can, in principle, be monitored by ERPs, they do not provide insights into the processing of complex prosodic information during speech perception.

The present paper gives an overview of six recent experiments which employed ERPs in order to shed some new light on the role of both prosody and punctuation during syntax parsing. Taking up Warren's (1999) suggestion, the experiments were designed to examine whether ERPs offer an appropriate way of studying prosodic processing online. In the following, I will give a brief introduction to the major issues relevant to all six experiments and then outline the basic design of the study.

Most previous studies on prosody and parsing have been inspired by Frazier's (1987) garden path model and focused on either early versus late closure (EC/LC) ambiguities or minimal versus non-minimal attachment (MA/NMA) ambiguities. Reading studies with unpunctuated sentences consistently reported disadvantages in EC and NMA structures as compared to LC and MA structures, as illustrated below in (1) and (2).

(1) LC *Since Jay always jogs <u>a mile and a half</u> this seems like a short distance to him*
(2) EC *Since Jay always jogs <u>a mile and a half</u> seems like a very short distance to him*

The sentence parser's LC preference to interpret the ambiguous noun phrase *a mile and a half* as the object of the *preceding* verb, leads to initial misunderstandings in the EC sentence (2) (Frazier & Rayner, 1982). However, in spoken language, this garden path effect was found to be diminished, apparently because a prosodic boundary after *jogs* in the EC sentence prevents the usual LC parsing preference (Marslen-Wilson et al., 1992; Speer et al., 1996; Warren et al., 1995). The observations of an immediate disambiguating influence of prosody on parsing decisions did not remain undisputed and others failed in replicating the findings (Watt & Murray, 1996). It was suggested that weaknesses of the experimental design might have caused the effects. In order to collect data, the cross-modal priming task usually employed required discontinuation of the speech presentation and subjects' performance in a quite unnatural task, thereby enhancing the risk of rather artifactual processing strategies. Whether phonological differences in the sentence materials led to inconsistencies could not be tested due to insufficient characterization of the speech signals. Authors who failed to replicate early effects argued in favor of a delayed rather than immediate influence of prosody.

In psycholinguistic reading studies, only few experiments investigated the role of punctuation, although additional studies were conducted most recently (Clifton, 1993; Cohen et al., 2001; Hill & Murray, 2000 and unpublished data; Mitchell & Holmes, 1985). In general, these studies support the view of a disambiguating potential of commas similar to that of prosodic speech boundaries. However, whether the comparable impact of commas and prosodic boundaries rests on a common mechanism, or rather on convergent but modality specific processes, remained largely speculative.

---

[1] The term 'emitted potential' refers to a biphasic ERP pattern that can be observed if, in a series of auditory stimuli, one stimulus is expected but not presented on time. That is, these potentials are emitted in absence of a physical stimulus.

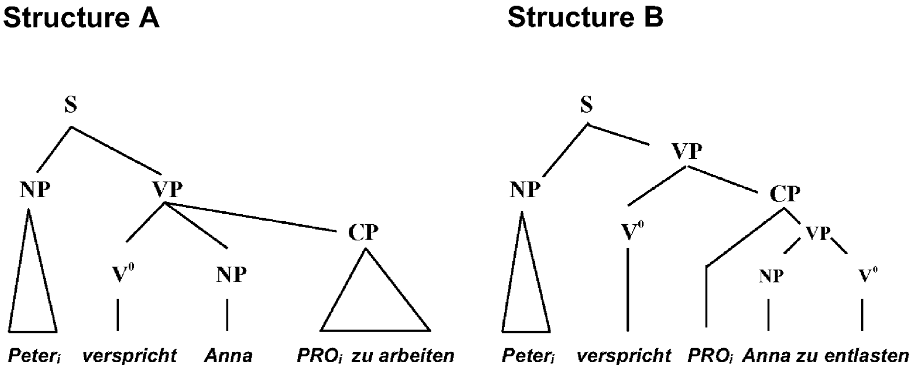

Fig. 1. Phrase markers of the two experimental sentence conditions A and B.

The present study adopts the EC/LC and MA/NMA approach of previous research and tries to meet the current standards of a solid acoustic characterization of the speech signals on which data interpretation can be based. Moreover, employing ERPs, it also tries to circumvent some of the methodological weaknesses of previous behavioral investigations. Finally, replicating the auditory experiments in a reading study should reveal whether the processing of commas could be related to the covert prosody of the inner voice. Supportive evidence may point to a direct link between punctuation and overt prosody in spoken language.

The stimulus material was derived from two German sentence types A and B which share the same word order but differ in their deep structure of syntactic relations (Fig. 1).

NP1 Verb1 NP2 Verb2
A. *Peter verspricht Anna zu arbeiten … und das Büro zu putzen*
(Peter promises Anna to work … and to clean the office.)
B. *Peter verspricht Anna zu entlasten … und das Büro zu putzen*
(Peter promises to support Anna … and to clean the office.)

Unlike in the English translation, word order information in German does not distinguish between A and B, and the structural differences are disambiguated lexically by the second verb alone. This Verb2 is intransitive in A (*arbeiten/work*) and obligatorily transitive in B (*entlasten/support*). Thus NP2 *Anna* is the indirect object of Verb1 (*verspricht/promises*) in A, but the direct object of Verb2 (*entlasten*) in B. According to the Minimal Attachment (MA) principle of the garden path model (Frazier, 1987), the ambiguous noun phrase *Anna* should initially be parsed as the object of the preceding verb. Thus, similar to the English examples in (1) and (2), A represents the preferred reading whereas in B this initial analysis must be revised and should result in a garden path effect when encountering the incompatible transitive Verb2 *entlasten*.[2]

The structural difference between A and B, particularly the deeper embedding of NP2 (*Anna*) in B, is expected to result in different prosodic patterns. Theories of syntax-prosody mapping (Selkirk, 1984) predict an additional prosodic boundary (#)

---

[2] Initially, sentence B should be erroneously perceived as *Peter promises Anna to support* which is a violation of the verb argument structure. Note that in German *das Büro/the office* cannot be interpreted as the object of *entlasten/support* unless it precedes the verb. Thus, the adequate German translation of the sentence *Peter promises Anna (both) to support and to clean the office* would be *Peter verspricht Anna das Büro zu entlasten und zu putzen.*

after Verb1 (*verspricht*) in B and a shift of the sentence accent from Verb2 (*arbeiten*) in A to NP2 (*Anna*) in B (words carrying accents are indicated by small capitals). Due to the additional boundary in B, this sentence consists of three intonation phrases (IPh), whereas A consists of only two IPh.

(A) *[**IPh1** Peter verspricht Anna zu ARBEITEN] # [**IPh2** und das Büro zu putzen]*
(B) *[**IPh1** Peter verspricht] # [**IPh2** ANNA zu entlasten] # [**IPh3** und das Büro zu putzen]*

The presence or absence of the first IPh boundary after Verb1 *verspricht* distinguishes *prosodically* between A and B much earlier than the *lexically* disambiguating information of Verb2 (underlined). Therefore it can be tested whether the prosodic information in B prevents the erroneous attachment of *Anna* to the preceding verb, which usually causes a garden path effect in reading studies. Moreover, if the early boundary after *verspricht* is introduced in sentence A, a prosodically guided parsing decision should result in a reversed garden path, i.e., the normally easy-to-process sentence A should become difficult. This reversed garden path sentence is illustrated in sentence example (C) and should be perceived as *Peter promises to work Anna ...* which in German is an outright violation of the obligatorily intransitive verb *arbeiten*.

(C) *[**IPh1** Peter verspricht] # [**IPh2** ANNA zu arbeiten] # [**IPh3** und das Büro zu putzen]*

In written German, the prosodic boundary after *verspricht* can be indicated by a comma, so that the stimulus material is compatible with both spoken and written language presentation.

The remainder of this paper is organized as follows. As the experiments shared many aspects in design and procedure, the subsequent Section 2 provides general information concerning subjects, materials and procedures that hold throughout the six experiments. (Further details can be found in Steinhauer, Alter, & Friederici, 1999; and Steinhauer & Friederici, 2001.) Sections 3–5 address several different aspects in turn, starting with acoustic and phonological characteristics of the speech signals and then turning to the behavioral and ERP findings of (1) the auditory and (2) the reading studies. One particular focus in Section 4 will lie on the hypothesized prosodic specificity of the closure positive shift (CPS) and its relation to other, already established, ERP components. Finally, Section 6 summarizes the most important claims and clarifies the differences between CPS and P600/SPS.

2. Materials and methods

2.1. Subjects

All 109 subjects were healthy, right-handed students of the University of Leipzig between 18 and 32 years of age. Each subject participated in one of the experiments only and was paid for participation.

2.2. Stimuli

Stimulus materials consisted of two main types of sentence constructions, hereafter referred to as (1) "experimental sentences" and (2) "filler sentences" although the latter were also systematically varied and analyzed in one of the reading experiments. The experimental sentences A, B, and C have already been characterized in the introduction. For the auditory experiments, sentence type C was derived from the speech signals of conditions A and B using the cross-splicing technique.

Filler sentences belong to two main types, F1 and F2:

NP1 Verb1 NP2 NP3 conj NP4
F1. *Der Mann sah die Frau, das Mädchen und den Großvater jeden Sommer im Urlaub*
(The man saw the woman, the girl and the grandfather every summer during the holidays)
NP1 Verb1 NP2 NP3 Verb2 NP4
F2. *Der Mann sah die Frau, das Mädchen sah den Großvater, und der Neffe sah die Tante*
(The man saw the woman, the girl saw the grandfather, and the nephew saw the aunt.)

The two filler sentences F1 and F2 are coordinative structures, i.e., they contain enumerations of equivalent syntactic constituents. F1 contains a series of three direct object NPs (NP2–NP4) of Verb1; F2 contains a series of three clauses with identical structure. A third filler condition (48 sentences) which is irrelevant for the present paper was derived by cross-splicing from conditions F1 and F2.

Of both experimental and filler sentences, 48 matched sentence pairs (A + B, F1 + F2) were constructed using frequency counts and phonological constraints according to the CELEX database. For the auditory experiments, all sentences were recorded with a trained female native speaker of German and stored as individual speech files (wav format) at a 44.1 kHz sampling rate and 16-bit resolution (CSL Model 4300B; Kay Instruments). Comprehension questions (e.g., *Does Anna promise to clean the office?*) were recorded with a male native speaker and stored in the same fashion. Written versions of the same sentences were used in the reading experiments.

Acoustic analyses and further processing of the speech files. The speech files were individually subjected to exhaustive acoustic analyses in order to determine the acoustic correlates of their prosodic structure in duration, pitch contour, and amplitude. Subsequently, they underwent further processing (i.e., cross-splicing, pause deletion, or filtering) in order to derive additional experimental conditions (see Sections 3–5).

2.3. Procedure

Subjects were seated in a dimly lit shielded chamber. One hundred and forty-four experimental and 144 filler sentences were presented in a pseudo-randomized order in blocks of typically 36 trials each, distributed across two sessions. Auditory stimuli were presented via two loudspeakers, visual stimuli in the center of a computer screen, controlled by ERTS software. In auditory experiments, each trial started with visual presentation of a fixation cross which remained on the screen while after 2 s the speech signal was presented for approximately 3.5 s. In Experiment 1, the display of a question mark in 20% of the trials indicated the subsequent auditory presentation of a 'yes/no' comprehension question that had to be answered by button press. Upon the response, exclamation marks replaced the question mark and indicated a time interval of 2500 ms during which subjects were encouraged to blink their eyes. In 80% of trials without questions, this eye blink interval started immediately after sentence presentation. After 2500 ms, the next trial started with the presentation of the fixation cross. In Experiments 2 and 3, subjects had to judge the prosodic acceptability immediately after *each* sentence ('good/bad' button press) and to answer subsequent comprehension questions in 20% of the trials. In reading experiments, sentences were presented word-by-word and as similar to the auditory experiments as possible. Thus, duration of display on the screen for each word position in the sentences was derived from mean word durations at this position in the speech signal, and commas attached to the preceding word mimicked the phrase boundary information of the speech signals. Each word was immediately replaced by the next word. While the comprehension task in 20% of the trials was adopted from the auditory experiments, an 'easy-to-read' judgment task ('easy/difficult' button press) replaced the prosodic judgment task. Further details will be given in the context of each experiment.

2.4. EEG recording and data analysis

EEG was recorded continuously from 17 cap-mounted tin electrodes (Electrocap International), referenced against the left mastoid. Impedances were kept below 5 kΩ. The midline comprised FZ, CZ, and PZ. Lateral electrodes were assigned to three (anterior, central, and posterior) regions of interest (ROIs) in each hemisphere (Steinhauer et al., 1999). Vertical and horizontal EOG was recorded bipolarly. EEG and EOG signals were amplified by a Neuroscan amplifier with built-in 40 Hz online low-pass filter and digitized at a sampling rate of 250 Hz and 12-bit resolution. One and 5 Hz low-pass filters were applied offline (see Sections 3–5).

EEG analyses were carried out with the EEP 3.0 software package (MPI of Cognitive Neuroscience). After artifact rejection, typically 35–40 trials (at least 30 trials) out of 48 trials per subject and condition entered the analyses. Percentage of rejections varied with epoch length but did not differ across conditions. Signal averages were calculated both across the entire sentence (0–4500 ms post sentence onset; mean = 35.4 trials) and for shorter 1000 ms intervals of relevant elements (e.g., Verb2; mean = 38.9 trials) using pre-stimulus baselines of 200 ms. In the auditory experiments, word onset times were identified via manually set markers in each individual speech signal. Amplitude and latency of ERP components were quantified by employing amplitude averages across representative time windows at all electrode sites, and by baseline-independent base-to-peak and peak-to-peak analyses, performed for each subject and each condition at midline electrodes.

2.5. Statistical analyses

The acoustic analyses (pitch values, word and pause durations, amplitude minima and maxima) included paired *t* tests and ANOVAs of the 48 speech files of conditions A and B. Performance data (judgments, error rates of the comprehension task, punctuation tests of the reading experiments) were analyzed conventionally with repeated measures ANOVAs after exclusion of outliers (>2 STD of the means). Analyses of variance of ERP data were generally performed separately for midline electrodes and the six lateral regions of interest (ROIs); the more time consuming base-to-peak and peak-to-peak analyses were restricted to midline electrodes. The design for midline electrodes included the within-subject factors *Condition* × *Electrode site* (3), that for lateral electrodes included the additional factor *Hemisphere*. Only significant effects involving at least one experimental factor will be reported. Potential violations of sphericity in designs with more than one degree of freedom in the numerator were addressed by the H-F correction procedure. To reduce the risk of progressive Type I errors, *p* values underwent a modified Bonferroni correction.

3. Results and discussion

3.1. Acoustic analyses of the speech signals

One intermediate step before the actual auditory experiments were run concerned the acoustic characterization of the speech files' prosodic structure. According to models of syntax-prosody mapping, the structural differences between the experimental sentences A and B should result in the presence of an additional boundary in B and the shift of the sentence accent position from Verb2 *arbeiten* in A to NP2 *Anna* in B. Both were confirmed by acoustic analyses. These analyses were performed on sentence fragments such as words and potential pause positions between words. Duration measures were collected for eight fragments via manually set markers in the speech signals (Cool Edit, Version 1.52). Analyses of the *F0* contour ("pitch tracking") considered minima and maxima at 11 representative positions in the speech signal (WinPitch, Version 1.8). The temporal loudness function (amplitude squares) was computed using scripts in MATLAB (Version 5.2) that accessed respective information for each sample point in the individual speech files.

The additional intonation phrase boundary in sentence B was realized most prominently by constituent lengthening of the first verb (*verspricht*) preceding the boundary and by a subsequent pause insertion as compared to sentence A. The sentence initial fragments comprising NP1 and Verb1 (e.g., *Peter verspricht*) had a mean duration of 781 ms in A and a significantly longer duration of 1179 ms in B [$F(1,47) = 793.92$; $p < 0.0001$]. Whereas in B there was a subsequent pause of 155 ms, virtually no pause (24 ms) was observed in A [$F(1,47) = 69.35$; $p < 0.0001$]. Except for the lexically different verbs, i.e., intransitive verbs such as *arbeiten* in A (802 ms) versus transitive verbs such as *entlasten* (717 ms) in B [$p < 0.0006$], no other elements displayed reliable duration differences between sentence types. Whether the longer verb duration in A has to be attributed to lexical or structural differences cannot be distinguished.

With respect to the sentence accent position, both pitch contour and amplitude measures indicated the predicted shift. In A the main pitch accent was found on Verb2 *arbeiten* (249.6 Hz) rather than on NP2 *Anna* (213.5 Hz), whereas in B the *F0* maximum occurred on NP2 *Anna* (262.4 Hz) rather than on Verb2 *entlasten* (234.1 Hz). This pattern was statistically confirmed by the highly significant interaction *Sentence type* × *Fragment* ($F(1,47) = 57.29$; $p < 0.0001$) and respective *Fragment* main effects in each sentence type (*p*s < 0.0001). Apart from these differences in the pitch contour, convergent evidence for dissimilar sentence accent positions in A and B was also derived from amplitude measures. The respective pitch accents were accompanied by local amplitude maxima, at Verb2 *arbeiten* in A, and at NP2 *Anna* in B. As with the pitch contours, this difference was reflected by a significant *Sentence type* × *Fragment* interaction [$F(1,47) = 39.26$; $p < 0.0001$].

Taken together, the acoustic analyses provided strong evidence for the expected prosodic differences between conditions A and B. As the prosodic boundary after *verspricht* in B was available earlier than the disambiguating *lexical* information of Verb2, this sentence material proved appropriate to investigate the potential real-time impact of prosodic information on parsing decisions. Importantly, such systematic prosodic differences were not only observed in the speech signals of the trained speaker. Recent production studies investigated the same sentence types and found converging evidence that healthy untrained speakers also reliably differentiate between conditions A and B (e.g., Schirmer, Alter, Kotz, & Friederici, 2001). Moreover, a study on 'reading prosody' demonstrated that subjects distinguished prosodically between structurally different sentences even when reading these aloud for the first time (Koriat, Greenberg, & Kreiner, 2002).

4. Auditory Experiments 1–3

4.1. Prosody-induced garden path effects

The first two auditory ERP experiments were conducted with 20 subjects each (Steinhauer et al., 1999). In Experiment 1 they had to answer comprehension questions only, in Experiment 2 they were asked to judge each sentence with respect to its prosodic appropriateness before answering comprehension questions. In both experiments it was expected that the early availability of disambiguating information, particularly the prosodic boundary after Verb1 *verspricht*, should prevent initial misunderstandings in the traditional garden path sentence B. Moreover, the incompatibility of this boundary with the intransitive verb *arbeiten* in mismatch condition C should cause a reverse garden path such that this usually easy-to-process sentence was initially perceived as **Peter promised to work Anna*, which is a verb argument structure violation in German.

In fact, all of these predictions were confirmed. As discussed elsewhere (Steinhauer et al., 1999), no indication of a traditional garden path was observed in condition B, whereas condition C displayed a strong reverse garden path effect in both behavioral and ERP data. The incompatible intransitive verb *arbeiten* in C elicited a biphasic pattern consisting of a central N400 and a subsequent parietal P600 as compared to the compatible transitive verb *entlasten* in B.[3] This finding is in line with reading studies on verb argument violations (Friederici & Frisch, 2000; Osterhout & Holcomb, 1992). Thus, prosodic information appears sufficient to not only prevent classical garden path effects (as in B) but to also cause strong reverse garden path effects (as in C). Additional peak-to-peak analyses including condition A ruled out the possibility that lexical verb differences between intransitive verbs (*arbeiten*) and transitive verbs (*entlasten*) contributed to the N400–P600 pattern, and rather confirmed that this pattern in condition C was due to a mismatch between the prosody-induced parsing and the intransitive verb's structural requirements. These ERP data provided the first evidence that prosody-induced (reversed) garden path effects are reflected by the same ERP components usually observed in lexically induced classical garden path sentences. Task differences between Experiments 1 and 2 did not seem to influence these effects.

4.2. The discovery of the CPS

Whereas the aforementioned garden path effects were predicted on the basis of previous ERP reading studies on verb argument violations, no prior work was available with respect to ERP correlates of prosodic processing as such. In the following, it will be demonstrated that the prosodic phrasing, which prevented the garden path effect in condition B and caused the reverse effect in C, reliably elicits a characteristic positive shift at speech boundaries, the CPS.

Fig. 2 shows a grand average ERP comparison of conditions A (upper panel) and B (lower panel) across the entire sentence at the PZ electrode. Each waveform

---

[3] Note that a direct comparison between conditions A and C (both of which contained the same verbs such as arbeiten) was not possible due to prosody-induced baseline differences (i.e., the CPS in B and C but not A; see Section 4.2).

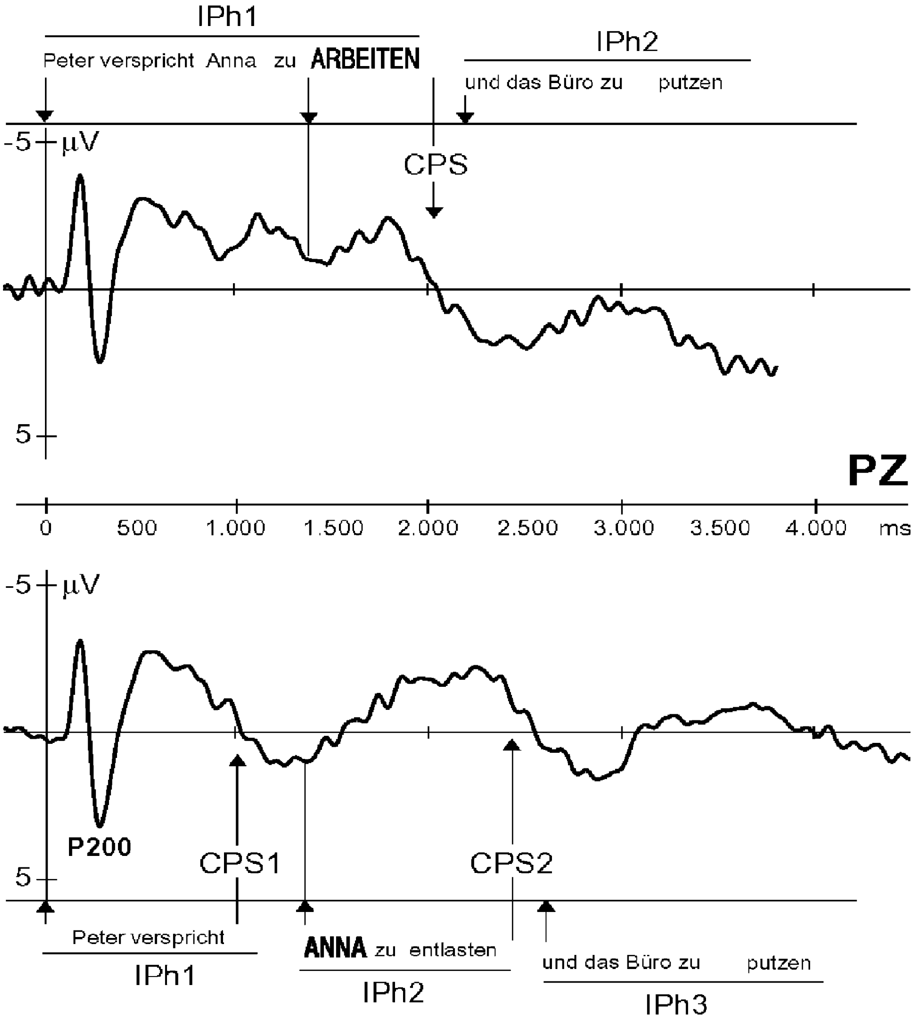

Fig. 2. The CPS in conditions A (upper panel) and B (lower panel) at PZ. The ERPs represent grand averages across all 40 subjects in Experiments 1 and 2. In both conditions, the CPS components occur exactly at the respective prosodic boundaries, in A at about 2000 ms, in B at 1000 and 2500 ms, respectively. As neither sentence condition contains any anomaly, the CPS components cannot be explained in terms of a P600. Note that negative amplitudes are generally plotted upward. (Modified from Steinhauer et al. (1999); Fig. 2, in *Nature Neuroscience, 2*, 191–196, with permission of Nature Publishing Group.)

comprises the data of all 40 subjects in Experiments 1 and 2 and is based on approximately 1500 single ERP signals.

In condition A, a single large positive shift can be observed at about 2000 ms after sentence onset, the latency of which coincides with the latency of the single intonation phrase boundary.[4] In condition B, two such positive shifts occur whose latencies also coincide with the latencies of the two phrase boundaries (at approximately 1000 and 2500 ms, respectively). The positive shifts were elicited in

---

[4] The offsets of verb1 (*verspricht*) and verb2 (*arbeiten*/*entlasten*), indicating the approximate position of prosodic boundaries in the speech signals, had the following average latencies: in A: 2161 ms (verb2); in B: 1179 ms (verb1) and 2573 ms (verb2). Although considerable latency variability across trials has to be taken into account, the somewhat shorter CPS onset latencies seem to suggest that the CPS was triggered while the verbs were still being processed.

absence of any violations or anomalies. This result suggests that the processing of prosodic boundaries results in a positive deflection in the listener's ERPs. As this novel component appears to be directly linked to the closure of prosodic phrases, it was termed the *closure positive shift* (CPS). If this interpretation holds, mismatch condition C with its two prosodic boundaries is expected to also display a pattern of two CPS components similar to condition B, whereas it should differ significantly from the lexically identical condition A. As illustrated by 1 Hz low-pass filtered ERPs in Fig. 3, this prediction was confirmed. Both conditions B and C show a common pattern of two CPS components and differ from the single CPS pattern in condition A. As a consequence, ERP amplitudes of condition A are more negative between 1000 and 1500 ms, and more positive between 2000 and 2500 ms than in conditions B and C. The superimposed ERPs of all three conditions also illustrate that the previously described biphasic N400–P600 garden path effect in condition C virtually 'rides' on the CPS of the second phrase boundary.

The CPS findings were statistically confirmed with a variety of approaches, employing both conventional amplitude averages and additional peak analyses. The latter were performed on the same 1 Hz low-pass filtered data that also entered the grand average in Fig. 3. Data in Table 1 demonstrate the statistical significance of

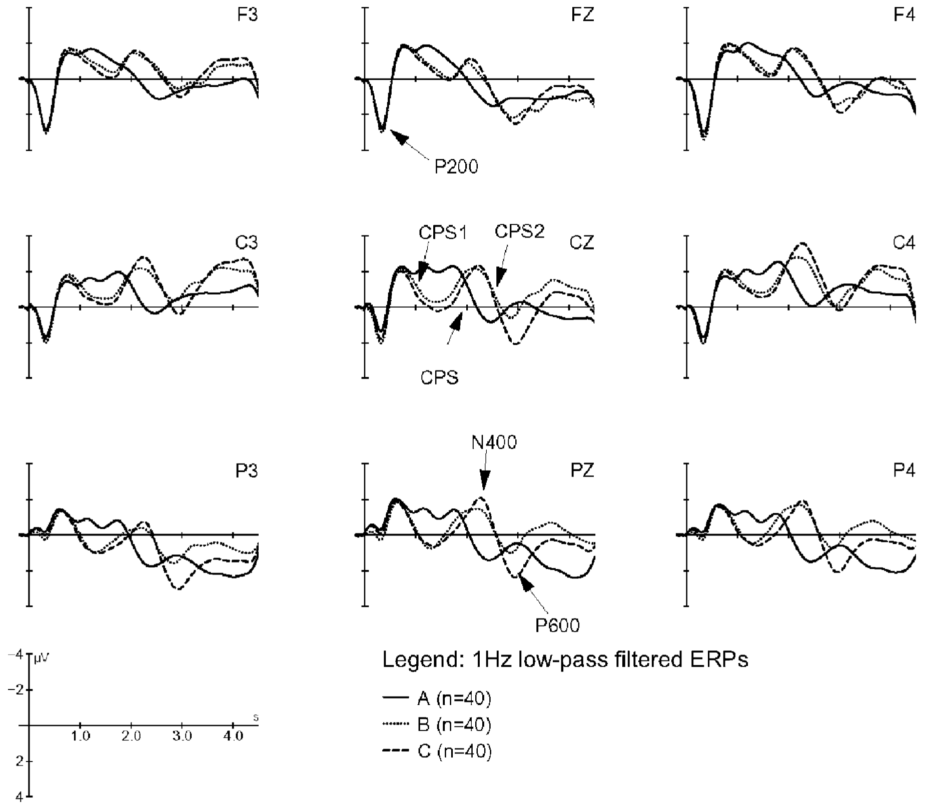

Fig. 3. Grand average ERPs across the sentence in all three conditions, displayed after 1 Hz low-pass filtering at nine electrodes. Condition A with its single CPS differs from both conditions B and C, which share the same pattern of two CPS components. At the disambiguating second verb, mismatch condition C elicits the N400–P600 pattern reflecting the reverse garden path effect. Whereas the amplitudes of the N100 and P200 components at sentence onset are considerably reduced (and virtually absent at PZ) due to the 1 Hz low-pass filtering, all of the CPS components remain unaffected (compared to Fig. 2). This indicates that phasic exogenous ERP components are unlikely to contribute to the CPS.

Table 1
CPS-effects per 500 ms time window at midline electrodes

	Time window (ms)	Source	F Value		p Value	ERP pattern[a]
			F(2/76)	F(4/152)		
1	0–500	—	—	—	—	—
2	500–1000	—	—	—	—	—
3	1000–1500	Cond.	9.07		.0003	A < B, C
		Cond. × Elec.		2.91	.0261	
4	1500–2000	Cond. × Elec.		2.68	.0419	
5	2000–2500	Cond.	9.59		.0002	A > B, C
		Cond. × Elec.		3.15	.0239	
6	2500–3000	—				
7	3000–3500	Cond. × Elec.		4.80	.0018	PZ : A, C > B
8	3500–4000	Cond. × Elec.		3.94	.0259	PZ : B < C

[a] ERP pattern X < Y: ERP amplitude of condition X is more negative than that of Y.

prosody-induced amplitude differences between condition A and the other two conditions in the time intervals of 1000–1500 ms and 2000–2500 ms at midline electrodes. Corresponding effects were found at lateral electrodes. Table 2 summarizes the somewhat more detailed amplitude and latency means in each condition based on peak analyses of amplitude minima preceding and amplitude maxima following potential positions of prosodic boundaries. Minima were quantified in three intervals: (1) 500–1650 ms, (2) 1650–2600 ms; (3) 2600–4000 ms. Maxima were measured in two intervals: (1) 500–2000 ms and (2) 2000–3500 ms. These intervals were chosen such that the entire time range was covered while overlaps of peaks with identical polarity across conditions were avoided. The amplitudes of the actual CPS components themselves were computed by peak-to-peak measures subtracting an amplitude minimum from the corresponding (i.e., subsequent) amplitude maximum. As shown in Table 3, condition A differed from both conditions B and C in its lack of the early CPS (after Verb1 *verspricht*) whereas the amplitude (although not the latency) of the second CPS (after Verb2) was identical between A and B but larger in C due to the additionally present N400–P600 garden path effect.

Table 2
Peak amplitudes and latencies in each time window at PZ

Condition		Time window (ms)	Peak	Latencies (ms)	Amplitudes (μV)	
	1	500–1650	min 1	900		−2.64
A		500–2000	MAX 1	1150		−0.15
	2	1650–2600	min 2	1900	CPS	−2.48
		2000–3500	MAX 2	2660		+2.10
	3	2600–4000	min 3	3150		−0.39
	1	500–1650	min 1	870	CPS 1	−2.17
B		500–2000	MAX 1	1340		+1.26
	2	1650–2600	min 2	2100	CPS 2	−2.21
		2000–3500	MAX 2	2840		+1.46
	3	2600–4000	min 3	3470		−1.27
	1	500–1650	min 1	750	CPS 1	−2.42
C		500–2000	MAX 1	1350		+1.48
	2	1650–2600	min 2	2210	CPS 2	−2.47
		2000–3500	MAX 2	2950		+2.7
	3	2600–4000	min 3	3570		−0.24

Note. min, minimum; MAX, maximum; CPS 1, first closure positive shift.

Table 3
Peak-to-peak effects for the first and the second positive shift at midline electrodes

Component	Condition contrast	Source	*df*[a]	*F* Value	*p* Value	ERP pattern[b]
	A:B:C	Cond	2.64	6.24	.0048	$A < B = C$
CPS 1	A:B	Cond	1.32	9.85	.0054	$A < B$
	A:C	Cond	1.34	9.95	.0051	$A < C$
	B:C	Cond	1.35	0.38	.8142	$B = C$
	A:B:C	Cond	2.34	5.20	.0161	$A = B < C$
CPS 2	A:B	Cond	1.15	1.93	.2768	$A = B$
	A:C	Cond	1.17	12.32	.0040	$A < C$
	B:C	Cond	1.29	10.76	.0040	$B < C$

[a] The degrees of freedom in the peak-to-peak analyses vary due to exclusion of subjects in each condition. In the ERPs of these subjects, the algorithm detecting the peaks did not find such local ERP maxima or minima, particularly at the FZ electrode. Additional analyses at single electrodes such as PZ increased the number of subjects considerably (average $n = 38$; range: 28–40), but did not change any of the results described here for the entire midline.

[b] ERP pattern $X < Y$: peak differences of condition X are smaller than those of condition Y.

4.3. Testing alternative accounts of the CPS

As the CPS was a new finding and, moreover, appeared to be the first prosody-related ERP component in the extant literature, particularly thorough analyses were required to rule out explanations in terms of other, already established ERP components. This section addresses the most important of these alternative accounts.

Contribution of exogenous components and sentence accents. First, it was possible that the CPS may result from exogenous components, such as the onset P200 of the critical second NP. Previous studies have shown that pauses in auditorily presented sentences enhance the N100–P200 complex of subsequent words (Besson et al., 1997; Holcomb & Neville, 1990). Given that conditions B and C contained a pause insertion at their first prosodic boundary, enhanced onset P200 components of the following NP2 *Anna* could have accumulated across trials and, due to latency variations, could have resulted in a CPS-like positive waveform. The following two arguments are to demonstrate that the CPS cannot be explained by exogenous components.

First, phasic components such as onset P200s were almost completely eliminated during 1 Hz low-pass filtering. As the CPS was robust in spite of the considerable reduction of the N100-P200 complex (compare Figs. 2 and 3 at PZ), a strong contribution of these components to the CPS is very unlikely. Second, as illustrated in Fig. 4, separate averages were computed for the critical NP2 (*Anna*) in all three conditions. Similar to previous studies, the preceding pause in conditions B and C leads to significantly enhanced onset P200s as compared to condition A without the pause (ps < 0.001). The obviously *frontal* maximum of the onset P200 rules out any direct contribution to the rather *posterior* CPS effects. Subsequent to the P200, condition A displays basically a similar pattern of negative and positive deflections as conditions B and C, however within a more positive amplitude range than these conditions. This pattern indicates (1) that further differences between condition A and the other conditions exist independent of the local P200 effects and (2) that these additional differences are likely to have occurred in the baseline interval of the current average (−200 to 0 ms), i.e., *before* the NP2 was presented. The negative offset in conditions B and C relative to A can be easily explained if their amplitudes were more *positive* during the baseline interval than those of condition A (as would be expected if the first CPS were present in this interval). The topographical distri-

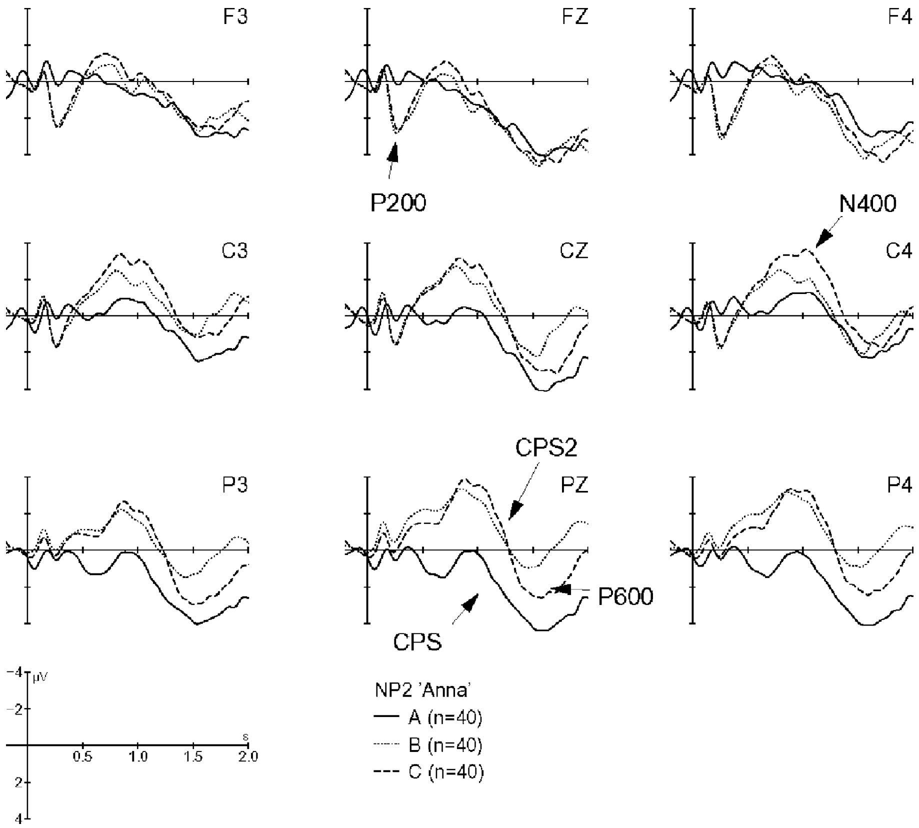

Fig. 4. Grand average ERPs for the second noun phrase (*Anna*). The frontal P200 pattern cannot account for the posterior CPS component. After 500 ms, conditions B and C display a negative offset relative to condition A, which is likely due to their first CPS being present in the baseline interval (Footnote 4). As a consequence of the temporal alignment to the onset of NP2, the other CPS after Verb2 occurs approximately at the same time in all three conditions. The presence of the CPS in conditions A and B illustrates that this component is independent of the P600, which occurs in condition C only.

bution of this offset with its posterior maximum and frontal minimum mirrors exactly that of the CPS components. These considerations strongly suggest that the negative offset in B and C is due to their first CPS being present in the baseline interval before the NP2 *Anna* was presented, which again rules out any contribution of NP2-related ERP components.

Importantly, this argument does not only hold for exogenous P200 components but also for any ERP effects associated with sentence accents. Recall that the sentence accents were carried by NP2 in conditions B and C, and by Verb2 in sentence A. Thus, both accents occurred too late in order to account for the presence of the first CPS.

Is the CPS a verb-related component? Positive going waveforms elicited by verbs were reported by Kutas (1997) and may reflect integration processes as verbs are viewed as the central element of a sentence. In the German sentence materials of the present study, prosodic phrase boundaries were generally preceded by a verb (*verspricht*; *arbeiten/entlasten*; *putzen*). Thus it was important to rule out the possibility that the CPS was actually identical to Kutas' verb-related positivity.

Two independent arguments can be advanced. First, conditions A and C were lexically identical, but a CPS after Verb1 was observed only in C. A relatively small positive deflection in condition A (Fig. 2) with a similar latency as the first CPS in B

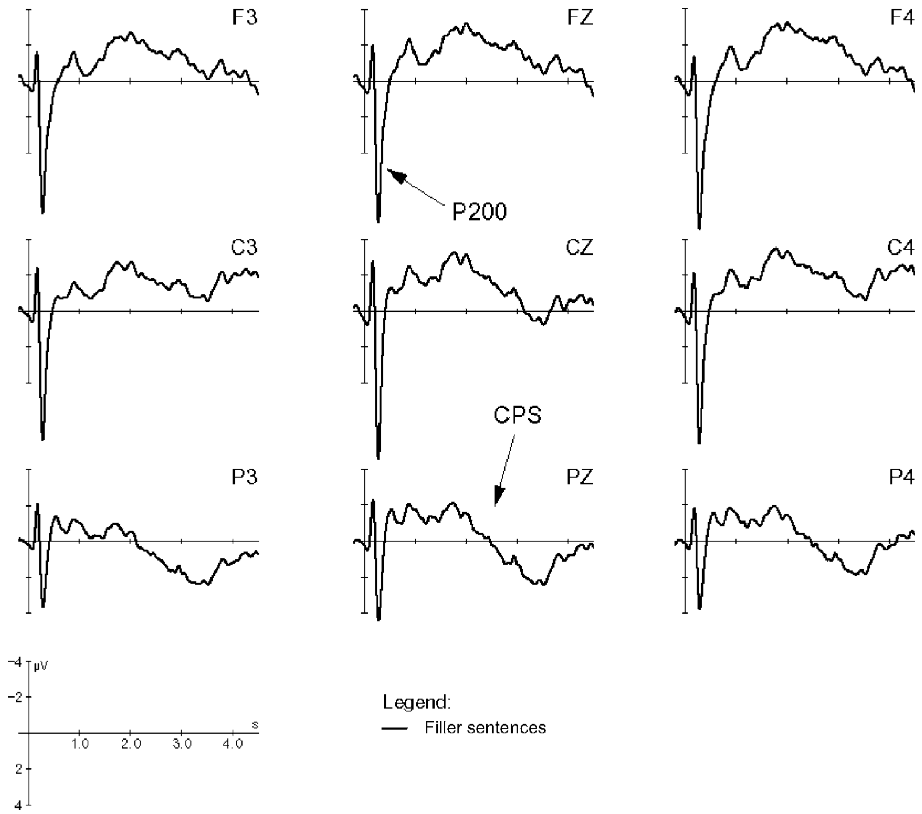

Fig. 5. Unlike with the experimental sentences, the CPS of the filler sentences was not elicited subsequent to a verb, but following a noun phrase. Thus, the component cannot be associated to verb specific processing.

and C, however, may indeed reflect verb-related processing. Second, unlike the experimental sentences, the filler sentences did contain prosodic boundaries that were *not* preceded by a verb but rather by a noun phrase. As Fig. 5 illustrates, this boundary also elicited a CPS similar to those in the experimental sentences.

4.4. Does the CPS reflect the processing of pauses?

The label ‘closure positive shift’ implicitly refers to the phonological (as opposed to the acoustic) level of prosodic processing in the listener. That is, the CPS is hypothesized to occur whenever the listener perceives a speech boundary, independent of whether it is acoustically realized by constituent lengthening, pause insertion, and/or boundary tones, all of which are known to be used by speakers to mark boundaries (so-called *cue trading*; e.g., Streeter, 1978). Alternatively, if the CPS can be shown to depend on one specific acoustic parameter (e.g., pause insertion), then its current interpretation has to be reconsidered.

Pause insertion was one of the most reliable acoustic boundary markers in both the sentence materials of the present study and the corresponding speech signals of untrained speakers (Schirmer et al., 2001). Intriguingly, in reading studies, sentence final words that are also usually followed by a pause have been reported to elicit P300-like positive ERP components as compared to mid-sentence words (Van Petten & Kutas, 1991). Thus, it was possible that the CPS was directly associated with the *acoustic* event of a pause rather than the *phonological* event of a boundary in the speech signal. In order to test this hypothesis, the pauses (i.e., intervals of silence) after the first verb *verspricht* in conditions B and C were carefully removed, resulting in the new conditions B′ and C′. Importantly, this procedure preserved small pause fragments (5–10 ms) in the signal in order to not affect the speech signals of adjacent words (e.g., *verspricht* and *Anna*, respectively). The duration of these fragments in B′ and C′, however, did not differ significantly from those in condition A without the boundary.

In the third ERP experiment, 16 subjects listened to these new conditions B′ and C′ (without the pauses) as well as to the original conditions A and B and to filler sentences (Steinhauer et al., 1999). In accordance with the cue-trading hypothesis, even after pause removal the remaining acoustic parameters (such as constituent lengthening) proved sufficient to mark the prosodic boundary in B′ and C′. Thus, subjects accepted B′ to the same extent as the original condition B, whereas the manipulated garden path condition C′ displayed a low acceptability rate of about 10%, similar to the old condition C in Experiment 2. In ERPs, the garden path effect in C′ was again reflected by an N400–P600 pattern. Most importantly, in contrast to condition A but similar to B, both new conditions B′ and C′ displayed a CPS component at their respective first boundary (Steinhauer et al., 1999). Thus, the CPS component is not associated with the presence of pauses in the speech signal. The data support the view that it reflects the processing of intonation boundaries as such.

5. Reading Experiments 4–6

5.1. Are commas equivalent to prosodic boundaries?

After the auditory experiments indicated that prosodic boundaries reliably elicit a CPS component in the listener's ERP, this component seemed to open a new approach for elucidating the 'inner voice' during silent reading. In particular, it could be tested whether commas in written sentences also elicit CPS-like components. This finding would provide the first direct evidence, that commas serve as visual triggers for covert prosodic phrasing.

The first ERP reading experiment (Experiment 4) was designed as a replication of the auditory experiments in the domain of written language (Steinhauer & Friederici, 2001). Sentence types A and B were presented with a comma (conditions 1a′ and 1b′) and without commas (conditions 1a and 1b) after the first verb *verspricht*. According to traditional German punctuation rules, sentence B requires a comma after *verspricht*, whereas A does not allow a comma. Thus sentences 1a and 1b′ meet the rules, whereas 1a′ and 1b do not. Condition 1b represents a classical garden path, whereas 1a′ represents a potential reversed garden path analogous to the mismatch condition C of the auditory experiments. During the ERP experiment, subjects were asked to read the sentences silently and to judge whether it was easy to read or not. In 20% of the trials they had to answer comprehension questions. Moreover, immediately after the ERP session subjects were presented with a list of unpunctuated sentences and had to insert commas according to their usual punctuation habits.

Fig. 6 shows the grand average ERPs of all 24 subjects across the entire sentence. In the two comma conditions 1a′ and 1b′, the first verb elicited a small but significant CPS-like positivity between 450 and 650 ms [main effect *Comma presence*: $F(2,22) = 4.84$; $p < 0.04$] which was followed by a large negative slow wave between 1050 and 2050 ms [$F(2,22) = 15.35$; $p < 0.0007$]. Interestingly, both components were significantly larger in a subgroup of 11 subjects who (1) inserted commas strictly according

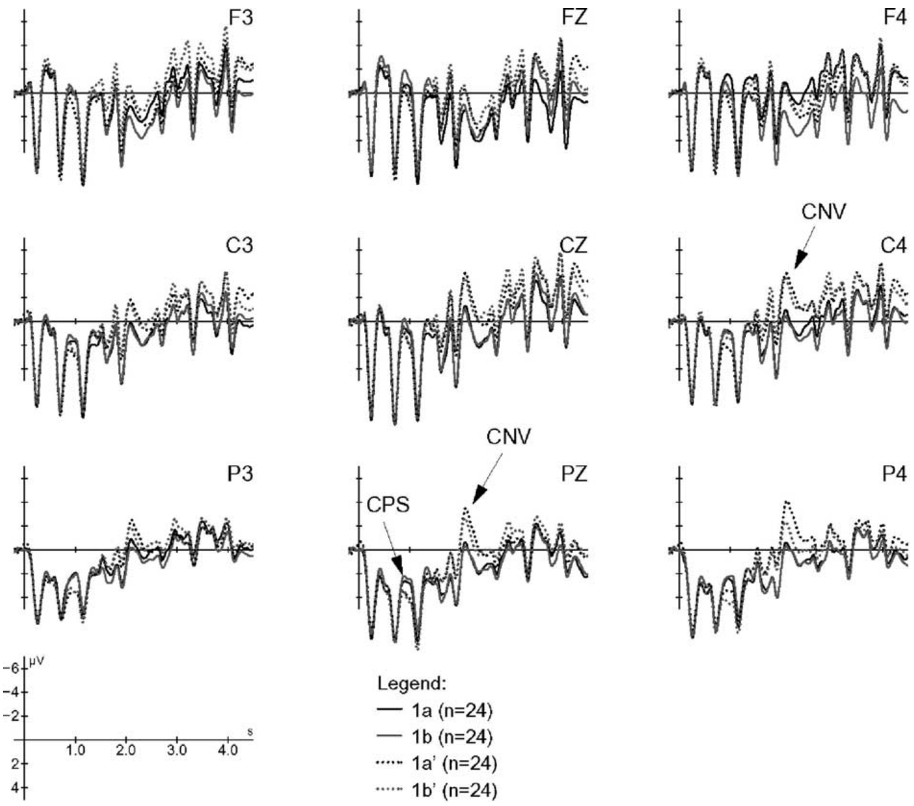

Fig. 6. ERPs of reading Experiment 4. CPS and negative slow wave elicited by comma conditions 1a′ and 1b′ in Experiment 4, shown as grand average ERPs across the entire sentence. Both components were larger in a subgroup of subjects with strict punctuation habits. While the slow wave was identified as a design-dependent CNV, the CPS proved robust across designs in Experiment 5 (see text).

to traditional German punctuation rules (no more than one error in the punctuation test), (2) were more susceptible to comma information when judging the reading difficulty of a sentence, (3) were guided by commas during parsing in a very similar way as were subjects of the auditory study by prosodic boundaries, and (4) performed better in the comprehension task as compared to the remaining 13 subjects (Steinhauer & Friederici, 2001). The latter group displayed inconsistent punctuation habits that violated traditional German rules (comma omissions: 51.9%; additional comma placement: 39.5%) and their easy-to-read judgments were virtually uninfluenced by punctuation. In ERPs, they did not show a significant CPS, and their slow wave was considerably smaller in amplitude (not illustrated). The group differences suggest an intra-individual correspondence of punctuation habits and the significance of comma information during reading. The ERPs seemed to reflect these differences.

In a follow-up experiment with 10 new subjects (Experiment 5), the negative slow wave could be identified as a design-dependent contingent negative variation (CNV; Tecce & Cattanach, 1987), likely reflecting expectancies concerning the potential reversed garden path (for details see Steinhauer & Friederici, 2001). The early CPS-like positive waveform, in contrast, proved a robust reflection of comma processing across experimental designs and could also be replicated at comma positions in filler sentences 2a′ and 2b′ [*Comma presence*: $F(1,9) = 8.63$; $p < 0.02$] (see sentence examples below and Fig. 7). Interestingly, the comma after NP3 (*das Mädchen/the girl*) is incompatible with traditional German punctuation rules in *both* types of filler

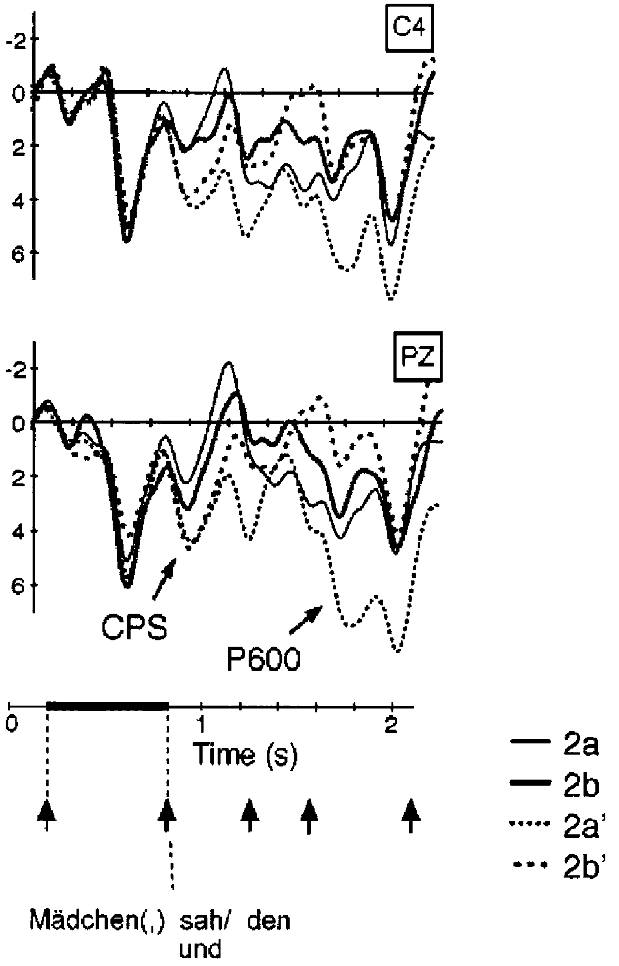

Fig. 7. In Experiment 5, commas in filler sentences 2a′ and 2b′ also elicited CPS components. In condition 2a′, the comma-induced boundary was not phonologically licensed and the disambiguating verb (*sah/saw*) elicited a subsequent P600. (Modified from Steinhauer & Friederici (2001) in the *Journal of Psycholinguistic Research, 30*, 267–295. Copyright Kluwer Academic Publishers. Reprinted with permission.)

sentences. Only in 2a′, however, does this comma also violate the phonological phrasing. The finding of a P600 in 2a′ but not 2b′ (Fig. 7) suggests that this phonological violation rather than the violation of punctuation rules as such causes processing difficulties.

(2a) *Der Mann sah den Jungen,* das Mädchen$_{(,)}$ *sah den Großvater, und...*
The man saw the boy, the girl$_{(,)}$ saw the grandfather, and...
(2b) *Der Mann sah den Jungen,* das Mädchen$_{(,)}$ *und den Großvater, während...*
The man saw the boy, the girl$_{(,)}$ and the grandfather, while...

As only the CPS-like positivity (but not the CNV component) was reliably elicited by commas, the findings in Experiments 4 and 5 support the idea that commas evoke ERP components similar to those elicited by prosodic boundaries.

5.2. Prosodic phrasing in reading without commas

The main difference between the CPS elicited by prosodic boundaries and that elicited by commas concerned their amplitude and duration. Both were larger in the auditory domain. It is conceivable that this difference might be accounted for by the lesser degree of activation of phonological representations during silent reading. That is, overt and covert prosody may differ primarily in ‘quantity’ rather than quality of phonological processing. If this hypothesis holds, then one would predict small CPS components for covert prosodic phrasing in general, no matter whether it is induced by commas or any other cue. This prediction was tested in the last experiment. In Experiment 6, sixteen subjects first listened to the de-lexicalized prosodic patterns of sentence types A and B of the auditory studies (‘prosody A’ and ‘prosody B’). These stimuli were derived by applying a particular filtering procedure to each speech signal in order to strip off all lexical (segmental) information while preserving the (suprasegmental) prosodic pattern with respect to pitch, amplitude and rhythm (for details see Meyer, Alter, Friederici, Lohmann, & von Cramon, 2002). Two seconds after auditory presentation of such a ‘sentence melody,’ a written sentence of either type A or B was presented word-by-word on the computer screen, in the same manner as in Experiments 4 and 5, however always without commas. Subjects were instructed to read the sentences silently and to replicate the previously heard sentence melody while reading. Thus unlike in Experiments 4 and 5, boundary information was provided by previous auditory input rather than punctuation. It was expected that the silent replication of the phrase boundary in prosody B during reading (covert prosody) should elicit a CPS similar to that induced by commas. Moreover, written sentences which were inconsistent with the previous prosodic pattern (e.g., prosody B followed by sentence A; B → A) should result in garden path effects. Finally, this design allowed one to test whether the processes underlying the CPS are syntactic rather than prosodic in nature. As syntactic processing crucially depends on lexical information, only a prosody-related CPS can be expected while subjects listen to de-lexicalized sentence melodies.

Fig. 8 illustrates that these predictions were confirmed. First, the finding of a CPS at the first boundary while subjects listened to prosody B (as compared to A;

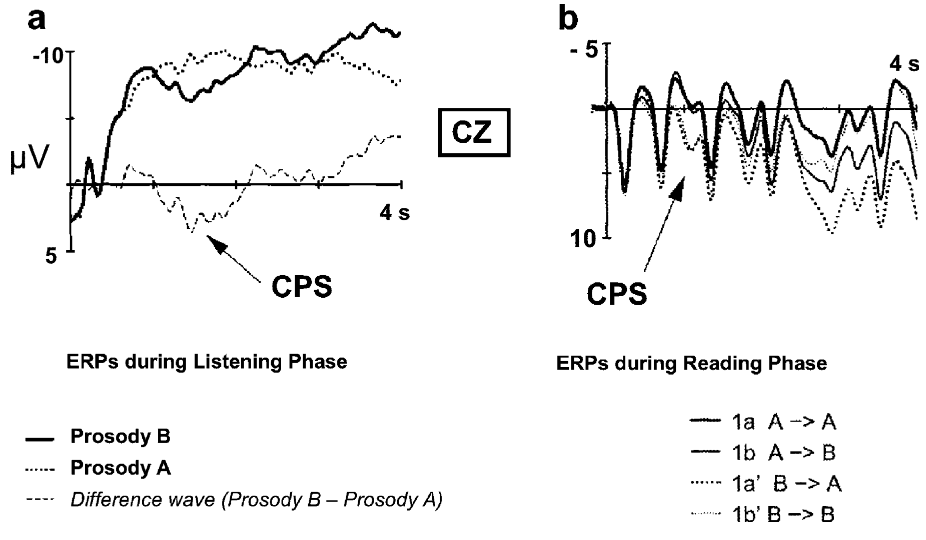

Fig. 8. ERP effects in Experiment 6 at CZ. (a) During the listening phase, de-lexicalized sentence melody B elicited a CPS component at its boundary. (b) After listening to sentence melody B, replication of its prosodic boundary during silent reading yielded a CPS similar to those elicited by commas (Fig. 6). Sentence structures incompatible with the prosodic pattern (A–B; B–A) also elicited P600 components at the disambiguating second verb. (Modified from Steinhauer & Friederici (2001) in the *Journal of Psycholinguistic Research*, *30*, 267–295. Copyright Kluwer Academic Publishers. Reprinted with permission.)

$p < 0.04$) provides strong evidence against a syntactic ERP component (Fig. 8a). Second, silent replication of prosody B, but not A, elicited a CPS at the corresponding boundary position, i.e. after the first verb (Fig. 8b; $p < 0.02$). The finding that silent replication of a prosodic boundary elicits the same ERP response as the commas in Experiments 4 and 5 (compare Figs. 6 and 8b) supports the view that commas serve as triggers for subvocal prosodic phrasing. Finally, the respective disambiguating verbs of sentences that were incompatible with the prosodic patterns (conditions $A \rightarrow B$ and $B \rightarrow A$) elicited a P600 effect as compared to compatible sentences ($p < 0.03$).

6. General discussion

In a series of six ERP experiments, it was shown that both prosodic boundaries in natural speech and commas in written language guide and even reverse syntactic parsing decisions immediately, and that these influences can be monitored online with ERPs. Moreover, the hypothesis was tested that their shared function of segmenting sentences into smaller phrases rests on the same mechanism of prosodic phrasing. ERP data support this hypothesis. Both prosodic boundaries and commas elicit the same brain response reflected by the CPS.

In auditory Experiments 1–3 it could be shown that this new ERP component was not due to other factors such as exogenous ERP components, the word class of the element preceding the boundary, the sentence accent, or the acoustic event of a pause in the speech signal. Moreover, the presence of a CPS for the de-lexicalized sentence melody in Experiment 6 indicates that this component cannot be linked to syntactic processing either. Therefore, the CPS component is taken to reflect the closure of prosodic phrases during language processing. The reliability of the auditory CPS effect has been demonstrated in several recent replications, including different sentence structures (U. Toepel, D. Saddy & K. Alter, unpublished data), cross-linguistic evidence (C. Brown & P. Hagoort, 2000, unpublished data), and data from children (K. Leuckefeld, A. Hahne & K. Alter, unpublished data). Yet another study showed that the CPS could be utilized to elucidate the relation between information structure, e.g. narrow focus, and prosodic phrasing (Hruska, Alter, Steinhauer, & Steube, 2000).

Experiments 4–6 examined the processing of commas in silent reading. Inspired by Chafe's (1988) suggestion of a direct correspondence between punctuation and prosody, it was tested whether commas can be viewed as orthographic triggers for covert prosodic phrasing. Employing the sentence materials of the auditory experiments in a reading experiment, with commas mimicking the prosodic boundaries, similar although smaller CPS components were found in both experimental and filler sentences. The replication of the comma-induced CPS component while subjects replicated a prosodic boundary during silent reading (Experiment 6) strongly supports the notion of a direct correspondence between punctuation and implicit prosody. The similarities point to a common mechanism and suggest that commas are in fact likely to trigger subvocal prosodic phrasing. Performance data revealed that subjects who use strict punctuation in writing are also more susceptible to commas during reading. Participants with less consistent punctuation habits were influenced by commas to a considerably lesser extent, performed worse in a comprehension task, and displayed no CPS component in the ERPs. Interestingly, accuracy in sentence comprehension was positively correlated ($p < 0.01$) with performance in the reading span test (Daneman & Carpenter, 1980). As the CPS amplitude and duration were generally reduced in reading (covert prosody) as compared to speech processing (overt prosody), the component seems to directly reflect the degree to which phonological representations are activated. Thus, the finding of a reduced CPS amplitude in poorer readers may indicate that they show an even more reduced activation of phonological representations as compared to skilled readers, which would be just the opposite pattern of what has previously been suggested (Bruthiaux, 1993). From this perspective, activation of phonological representations during reading (i.e., listening to the ‘inner voice’) appears to support sentence comprehension.

6.1. Profile and functional significance of the CPS

With the CPS, a new ERP component in addition to LAN, N400 and P600/SPS has become available to psycholinguistic research. Unlike the latter components, it does not primarily reflect semantic or syntactic processing but phonological/prosodic phrasing, thus opening a window to another domain of linguistic online processing. Its main characteristics can be summarized as follows. The CPS is a bilateral, centro-parietal positive deflection, most prominent at midline electrodes. Largely independent of input modality, the task employed, and the type of cue (acoustic parameters, commas), it seems to be elicited online whenever listeners or readers segment the current sentence into prosodic phrases. The CPS amplitude is larger in auditory (Experiments 1–3) than visual language presentation (Experiments 4 and 5) and appears to correlate positively with the degree to which phonological representations are activated. In listeners the mean amplitude of the shift varied between 3.4 and 4.6 μV, with durations of some 500 ms (Table 3), while readers showed reduced CPS amplitudes (approximately 2 μV) and shorter durations (200–400 ms). The component's onset at speech boundaries was likely triggered by the first available acoustic boundary marker, possibly pre-final constituent lengthening. This interpretation would explain why the CPS onset preceded the onset of the pause (Footnote 4) and why the component still occurred after pause removal (Experiment 3). However, taking into account acoustic variability across the presented speech files and potential differences in phonological susceptibility among listeners, only future research will be able to reveal the exact acoustic and phonological conditions/circumstances under which the CPS is reliably elicited. Only after the critical events triggering the ERP component have been identified will it be possible to determine the actual onset latency of the auditory CPS in more detail. During silent reading, CPS components were observed with latencies of some 400 ms after onset of the word to which the comma was attached.

One of the most important questions concerning the CPS is whether the strong claim of a novel ERP component is justified, or whether it can rather be explained in terms of other, already established ERP components. Whereas the contribution of verb-related positivities (Kutas, 1997) or exogenous components such as the P200 could not account for the observed effects (Fig. 4), the relationship between CPS and P600/SPS still needs to be addressed. This issue deserves particular attention as both CPS and P600/SPS share a number of characteristics that might suggest that they may be the same component. Both ERP effects are associated with language processing, have a positive polarity, and display a bilateral centro-parietal scalp distribution, with largest amplitudes at posterior electrodes. Therefore, it seems important to emphasize the data that support a distinction between CPS and P600.

First, whereas the P600/SPS is commonly viewed as a reflection of additional *syntactic processing* (particularly in case of syntactic anomalies; Hagoort et al., 1993; Hahne & Friederici, 1999; Osterhout & Holcomb, 1992), the CPS is assumed to reflect *prosodic phrasing*. In fact, the sentence positions where the CPS was observed in the auditory experiments did not coincide with positions of syntactic anomalies but rather with positions of prosodic boundaries. Neither the corresponding ERP differences between sentences A and B in Fig. 2 nor the CPS found in filler sentences (Fig. 5) involved any anomalies. The same holds for the reading experiments. The CPS in the reading phase of Experiment 6 (Fig. 8b) in particular supports the notion of prosodic phrasing underlying this component. Second, CPS and P600 behave additively. In mismatch condition C of Experiments 1 and 2, the *first* boundary (CPS 1) led to a syntactic anomaly at the subsequent disambiguating Verb2 (*arbeiten*), resulting in the P600 at the *second* boundary (CPS 2). The amplitude of this garden path effect P600 adds to the CPS of the second prosodic boundary (Figs. 3 and 4), suggesting independent neural generators of the two components. It is also crucial to recall that the other four CPS components (Tables 2 and 3) were quantified in complete absence of any N400s or P600s. Third, during the initial listening phase of Experiment 6, a CPS was elicited by a speech boundary in de-lexicalized sentence melodies. As lexical information is indispensable for syntactic processing (including the closure or 'wrap-up' of *clauses* which rather seem to elicit *negative* ERP components; Kutas, 1997), this finding is incompatible with a syntactic CPS account. Finally, both the longer duration and the larger amplitude of CPS components at speech boundaries as compared to comma positions point to a phonological effect and are difficult to account for with a purely syntactic CPS interpretation. As a whole, only the CPS interpretation in terms of a reflection of prosodic phrasing is compatible with all data points, including those of recent CPS replications.

References

Bader, M. (1998). Prosodic influences on reading syntactically ambiguous sentences. In J. D. Fodor & F. Ferreira (Eds.), *Reanalysis in sentence processing* (pp. 1–46). Dordrecht: Kluwer.

Besson, M., Faita, F., Czternasty, C., & Kutas, M. (1997). What's in a pause: Event-related potential analysis of temporal disruptions in written and spoken sentences. *Biological Psychology, 46*, 3–23.

Böcker, K. B. E., Bastiaansen, M. C. M., Vroomen, J., Brunia, C. H. M., & Gelder, B. (1999). An ERP correlate of metrical stress in spoken word recognition. *Psychophysiology, 36*, 706–720.

Bruthiaux, P. (1993). Knowing when to stop: Investigating the nature of punctuation. *Language and Communication, 13*, 27–43.

Chafe, W. (1988). Punctuation and the prosody of written language. *Written Communication, 5*, 396–426.

Chwilla, D. J., Brown, C. M., & Hagoort, P. (1995). The N400 as a function of the level of processing. *Psychophysiology, 32*, 274–285.

Clifton, C., Jr. (1993). Thematic roles in sentence parsing. *Canadian Journal of Experimental Psychology, 47*, 222–246.

Cohen, H., Douaire, J., & Elsabbagh, M. (2001). The role of prosody in discourse processing. *Brain and Cognition, 46*(1–2), 73–82.

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.

Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech, 40*, 141–201.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450–466.

Fodor, J. D. (1998). Learning to parse. *Journal of Psycholinguistic Research, 27*, 285–319.

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and Performance XII* (pp. 559–586). Hillsdale, NJ: Erlbaum.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14*, 178–210.

Friederici, A. D., & Frisch, S. (2000). Verb-argument structure processing: The role of verb-specific and argument-specific information. *Journal of Memory and Language, 43*, 476–507.

Friederici, A. D., Mecklinger, A., Spencer, K. M., Steinhauer, K., & Donchin, E. (2001). Syntactic parsing preferences and their on-line revisions: A spatio-temporal analysis of event-related brain potentials. *Cognitive Brain Research, 11*(2), 305–323.

Friederici, A. D., Steinhauer, K., & Frisch, S. (1999). Lexical integration: Sequential effects of syntactic and semantic information. *Memory & Cognition, 27*, 438–453.

Friederici, A. D., Steinhauer, K., Mecklinger, A., & Meyer, M. (1998). Working memory constraints on syntactic ambiguity resolution as revealed by electrical brain responses. *Biological Psychology, 47*, 193–221.

Friederici, A. D., Steinhauer, K., & Pfeifer, E. (2002). Brain signatures of artificial language processing: Evidence challenging the critical period hypothesis. *Proceedings of the National Academy of Sciences USA, 99*, 529–534.

Gunter, T. C., Stowe, L. A., & Mulder, G. (1997). When syntax meets semantics? *Psychophysiology, 34*, 660–676.

Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift as an ERP-measure of syntactic processing. *Language and Cognitive Processes, 8*, 439–483.

Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *Journal of Cognitive Neuroscience, 11*, 194–205.

Hill, R., & Murray, W. S. (2000). Commas and spaces: Effects of punctuation on eye movements and sentence parsing. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 565–590). Oxford: Elsevier.

Holcomb, P. J., & Neville, H. J. (1990). Semantic priming in visual and auditory lexical decision: A between modality comparison. *Language and Cognitive Processes, 5*, 281–312.

Hruska, C., Alter, K., Steinhauer, K., & Steube, A. (2000). Can wrong prosodic information be mistaken by the brain? *Journal of Cognitive Neuroscience* (Suppl. 122).

Koriat, A., Greenberg, S. N., & Kreiner, H. (2002). The extraction of structure during reading: Evidence from reading prosody. *Memory and Cognition, 30*, 270–280.

Kutas, M. (1997). Views on how the electrical activity that the brain generates reflects the functions of different language structures. *Psychophysiology, 34*, 383–398.

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

Marslen-Wilson, W. S., Tyler, L. K., Warren, P., Genier, P., & Lee, C. S. (1992). Prosodic effects in minimal attachment. *The Quarterly Journal of Experimental Psychology, 45*(1), 73–87.

Meyer, M., Alter, K., Friederici, A. D., Lohmann, G., & von Cramon, D. Y. (2002). FMRI reveals brain regions mediating slow prosodic modulations in spoken sentences. *Human Brain Mapping, 17*, 73–88.

Mitchell, D. C., & Holmes, V. M. (1985). The role of specific information about the verb in parsing sentences with local structural ambiguity. *Journal of Memory and Language, 24*, 542–559.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31*, 785–804.

Schirmer, A., Alter, K., Kotz, S. A., & Friederici, A. D. (2001). Lateralization of prosody during language production: A lesion study. *Brain and Language, 76*, 1–17, doi:10.1006/brln.2000.2381.

Selkirk, E. (1984). *Phonology and Syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.

Share, D. L. (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology, 72*, 95–129.

Speer, S. R., Kjelgaard, M. M., & Dobbroth, K. M. (1996). The influence of prosodic structure on the resolution of temporary syntactic closure ambiguities. *Journal of Psycholinguistic Research, 25*, 247–268.

Steinhauer, K., Alter, K., & Friederici, A. D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience, 2*(2), 191–196.

Steinhauer, K., & Friederici, A. D. (2001). Prosodic boundaries, comma rules, and brain responses: The closure positive shift in ERPs as a universal marker for prosodic phrasing in listeners and readers. *Journal of Psycholinguistic Research, 30*(3), 267–295.

Streeter, L. A. (1978). Acoustic determinants of phrase boundary location. *Journal of the Acoustical Society of America, 64*, 1582–1592.

Tecce, J. J., & Cattanach, L. (1987). Contingent negative variation (CNV). In E. Niedermeyer & F. Lopes da Silva (Eds.), *Electro-encephalography: Basic principles, clinical applications and related fields* (pp. 658–679). Munich: Urban & Schwarzenberg.

Van Petten, C., & Kutas, M. (1991). Influences of semantic and syntactic context on open- and closed-class words. *Memory & Cognition, 19*, 95–112.

Warren, P. (1999). Prosody and language processing. In S. Garrod & M. Pickering (Eds.), *Language Processing* (pp. 155–188). Hove: Psychology Press.

Warren, P., Grabe, E., & Nolan, F. (1995). Prosody, phonology, and parsing in closure ambiguities. *Language and Cognitive Processes, 10*, 457–486.

Watt, S. M., & Murray, W. S. (1996). Prosodic form and parsing commitment. *Journal of Psycholinguistic Research, 25*, 291–318.