

Effects of Cooperating and Conflicting Prosody in Spoken English Garden Path Sentences: ERP Evidence for the Boundary Deletion Hypothesis

Efrat Pauker, Inbal Itzhak, Shari R. Baum, and Karsten Steinhauer

Abstract

■ In reading, a comma in the wrong place can cause more severe misunderstandings than the lack of a required comma. Here, we used ERPs to demonstrate that a similar effect holds for prosodic boundaries in spoken language. Participants judged the acceptability of temporarily ambiguous English “garden path” sentences whose prosodic boundaries were either in line or in conflict with the actual syntactic structure. Sentences with incongruent boundaries were accepted less than those with missing boundaries and elicited a stronger on-line brain response in ERPs (N400/P600 components). Our results support the notion that mentally deleting an overt prosodic boundary is more

costly than postulating a new one and extend previous findings, suggesting an immediate role of prosody in sentence comprehension. Importantly, our study also provides new details on the profile and temporal dynamics of the closure positive shift (CPS), an ERP component assumed to reflect prosodic phrasing in speech and music in real time. We show that the CPS is reliably elicited at the onset of prosodic boundaries in English sentences and is preceded by negative components. Its early onset distinguishes the speech CPS in adults both from prosodic ERP correlates in infants and from the “music CPS” previously reported for trained musicians. ■

INTRODUCTION

In language, ambiguities are abundant at most levels of linguistic analysis. In addition to multiple meanings at the word level, such as homophones (e.g., *I* vs. *EYE*) and homonyms (such as *World BANK* vs. *river BANK*), temporary uncertainty also occurs at the sentence level due to ambiguous syntactic structures. In many cases, these ambiguities are successfully resolved by our brain within a few hundred milliseconds without even reaching awareness (Frazier, 1987). However, in cases when the human sentence processor initially commits to the wrong interpretation, such ambiguities can cause severe misunderstandings called “garden path” effects. For example, the sentence, *Mary said Peter’s brother is a nice girl*, is likely to elicit a surprise response and additional processing unless prosodic boundaries¹ or commas before and after “said Peter’s brother” clarify the intended parenthetical structure. In the absence of such cues, the *conceptual–semantic* implausibility of the initial sentence interpretation may lead to a re-evaluation of the structure (e.g., by re-reading the sentence)—and, ultimately, to its resolution. In other cases, lexical information further downstream may provide *syntactic* cues signaling the initial mistake. One of the best-studied garden path structures of this kind is the classical early closure (EC) versus late closure (LC) ambiguity (Frazier

& Rayner, 1982). Consider the following ambiguous sentence fragment in (1):

(1) *Whenever John walks the dog ...*

(1a) *... the kids are chasing him.* (1b) *... is chasing him.*

The verb “walks” is optionally transitive. As illustrated in Figure 1, it can either take a direct object—in this case, “the dog” [see Structure (1a)]—or it need not be followed by an object, as in Structure (1b), in which the noun phrase (NP) “the dog” represents the subject of a new clause. Importantly, the actual underlying structure remains ambiguous until lexical material following “the dog” has been encountered (*the kids* vs. *is chasing*). Reading studies have consistently demonstrated a strong preference to initially interpret the NP “the dog” as the direct object of *walks*, thus supporting Structure (1a), and resulting in garden path effects for Sentence (1b) (e.g., Frazier & Rayner, 1982).

Different models have been proposed to account for how cases of structural ambiguity like the one presented above are processed. First, various syntactic accounts suggest that the human sentence parser follows certain principles that initially favor the simplest possible syntactic structure compatible with the input (Gorrell, 1995). In particular, Frazier (1987) and Frazier and Fodor (1978) proposed the garden path model, in which the parser relies on strategic guessing while making initial structural decisions, because such a strategy reduces the load on limited

McGill University, Montreal, Canada

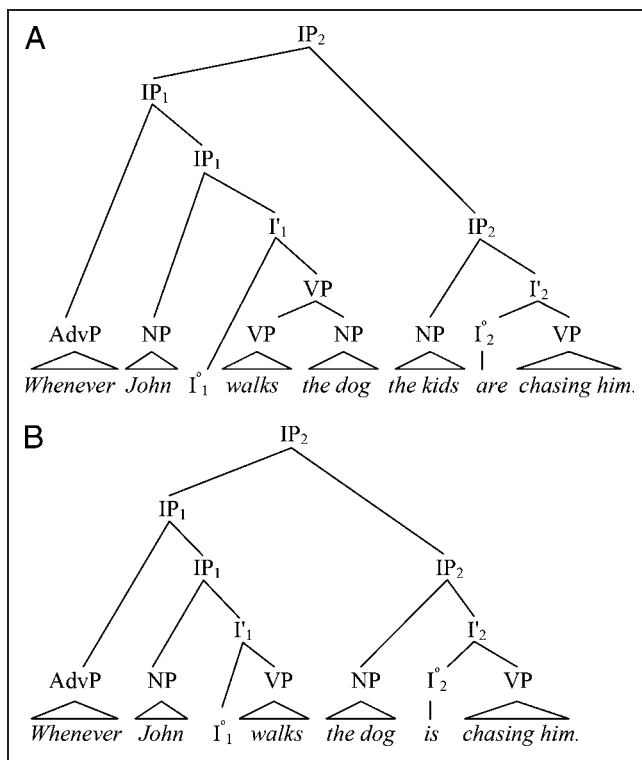


Figure 1. Phrase markers for (A) the preferred “transitive” late closure and (B) the nonpreferred “intransitive” early closure structures compatible with sentence fragment (1). AdvP = adverbial phrase; I(P) = inflection (phrase); NP = noun phrase; VP = verb phrase.

working memory capacity. If subsequent input confirms the initial interpretation, the ambiguity may not even be realized, whereas in cases of conflicting information further downstream, this type of approach may lead one up a “garden path.” Most relevant to the present investigation, their parsing principle of late closure (LC) states that “when possible attach incoming lexical items into the clause or phrase currently being processed” (Frazier & Fodor, 1978). This principle predicts that the parser will initially attempt to attach the ambiguous NP “the dog” under the verb phrase (VP) preceding it, which is compatible with (1a) but causes the garden path effect in (1b). Secondly, syntactic parsing preferences do not operate in isolation; they work in addition to, or in interaction with, other influencing factors. For instance, in EC/LC closure ambiguities, past tense forms (e.g., *Whenever he walked the dog...*) appear to strengthen the LC preference and to increase EC garden path effects, whereas progressive verb forms (e.g., *Whenever he was walking the dog...*) do not (Frazier, Carminati, Cook, Majewski, & Rayner, 2006). Similarly, a lexical bias of individual (optionally transitive) verbs toward transitive use (e.g., *fold*) also seems to support the LC preference compared to verbs with an intransitivity bias (e.g., *swim*) (Itzhak, Pauker, Drury, Baum, & Steinhauer, 2010; Staub, 2007). With another type of structural ambiguity, it has been shown that verb complement biases may lead to parsing decisions in the opposite direction than

predicted by syntax-based parsing principles (Wilson & Garnsey, 2009).

Garden Path Effects and the Role of Prosody and Punctuation

Although many reading studies support a preference for one structure over another (e.g., Frazier & Rayner, 1982), in keeping with Frazier and Fodor’s (1978) parsing principles, studies with auditory stimuli suggest that parsing choices are influenced by information that is absent in written text, namely, prosodic information. Prosody refers to the intonation and rhythmic pattern of spoken sentences, including the presence of prosodic boundaries (e.g., pauses) that typically coincide with major syntactic boundaries. It has been demonstrated that speakers produce ambiguous sentences with prosodic cues to the intended interpretation, and that listeners make use of these cues in comprehension (Snedeker & Trueswell, 2003; Schafer, Speer, Warren, & White, 2000; Price, Ostendorf, Shattuck-Hufnagel, & Fong, 1991). In sentences containing temporary structural ambiguities, prosody often seems to serve a disambiguating role *before* disambiguating lexical information is encountered (Schafer et al., 2000; Marslen-Wilson, Tyler, Warren, Grenier, & Lee, 1992; Beach, 1991). An early prosodic boundary (#) after *walks* should thus prevent the otherwise preferred LC reading and, therefore, avoid the EC garden path effect in (1b): *Whenever John walks (#) the dog is chasing him*, as demonstrated by Marslen-Wilson et al. (1992) using a cross-modal naming paradigm.

In an influential study, Kjelgaard and Speer (1999) presented subjects with EC and LC sentences in three different conditions, varying the degree to which the prosodic pattern and the syntactic structure matched each other. The first condition demonstrated a cooperative relation between prosody and syntax, using well-formed sentences in which syntactic and prosodic boundaries coincided. The prosody in the second condition, labeled “baseline,” contained rather weak boundaries at both positions and was created to be equally compatible with both EC and LC structures. In the third condition, strong prosodic boundaries occurred at misleading locations, creating a mismatch between syntactic and prosodic boundaries. The level of processing difficulty was measured by response time, accuracy of speeded grammaticality judgments, metalinguistic judgments of comprehension, and sentence completion tasks. Across all tasks, EC and LC sentences in Kjelgaard and Speer’s study were equally easy to process when presented in the cooperating condition, convincingly demonstrating that the presence of prosodic cues overrode any parsing challenges for EC constructions that emerge in the absence of prosody, that is, while reading (Frazier & Rayner, 1982). Furthermore, the results show that conflicting prosodic boundaries can lead to garden path effects in both LC and EC sentences, indicating that speech prosody has a central role in guiding the parser to construct the initial syntactic representations.² Importantly, EC sentences

in this study were found to be significantly more difficult to process than LC sentences in both conflicting and baseline conditions. In those tasks providing the best on-line measure (i.e., cross-modal naming; Experiments 3 and 4), data indicated that LC sentences with conflicting prosody were indeed easier to process than EC sentences in the baseline condition. The overall pattern led the authors to suggest that the structural LC preference over EC may also play a strong role in auditory sentence processing and help listeners resolve initial misunderstandings.

However, not all findings have been consistent with this view. For instance, Walker, Fongemie, and Daigle (2001) confirmed the decisive role of prosodic boundaries in spoken EC/LC sentences, but did not find evidence for an overall advantage for LC structures. Similarly, Steinhauer and colleagues tested EC/LC ambiguities in German, employing both spoken (Steinhauer, Alter, & Friederici, 1999) and written sentences (Steinhauer & Friederici, 2001), and observed particularly strong garden path effects in LC sentences.³ The latter reading study used commas that, according to the authors, can trigger the subvocal generation of prosodic boundaries during silent reading. Thus, even though written text per se does not provide prosodic cues, punctuation may indirectly have very similar effects as overt prosodic boundaries in speech. Contrary to Kjelgaard and Speer's (1999) findings, their comma-induced LC garden path sentence turned out to be much more difficult than the "classical" EC garden path condition. One major factor that Steinhauer and Friederici (2001) proposed as underlying the processing difficulties was the type of prosodic pattern that conflicted with the syntactic structure. Their difficult LC garden path contained a comma that needed to be ignored (or to be "mentally deleted") in order to resolve the problem, whereas the easy "classical" EC garden path required readers to mentally insert a comma (or boundary) that was missing in the original sentence. The authors concluded that "the mental deletion of a previously assumed pause/comma/boundary may be more costly than the postponed insertion of an initially omitted pause/comma/boundary" (Steinhauer & Friederici, 2001). We will henceforth refer to this hypothesis as the "boundary deletion hypothesis" (BDH). However, given the modality differences between studies, it is conceivable that (contrary to Steinhauer and Friederici's claims) the disambiguating function of commas and prosodic boundaries does not rest on the same mechanisms. In other words, the BDH may hold for commas, but not for speech boundaries. A recent electrophysiological study by Kerkhofs, Vonk, Schriefers, and Chwilla (2008) seems, indeed, to cast doubt on such a direct correspondence between commas and prosodic boundaries as they found different ERP patterns in response to written versus spoken materials. In particular, a CPS (see below) was elicited by prosodic boundaries but not by commas. However, Liu, Wang, and Jin (2010) found CPS-like positivities for commas in Chinese. The variability in findings may thus be partly influenced by both cross-linguistic and interindividual differences in the

use of punctuation rules (see also Steinhauer & Friederici, 2001).

Event-related Brain Potentials and the Role of Prosody in Sentence Processing

Although the findings reviewed above are compelling, the majority of studies relied primarily on behavioral measures and subjective judgments. Using ERPs allows a more objective on-line means to investigate the interaction of prosody with syntactic parsing. Moreover, ERPs reflect auditory processing across the entire length of a sentence, and the profile of ERP components provides more specific information about the actual nature of processing difficulties. Several ERP components have been previously described in the literature in relation to sentence processing, a number of which are particularly relevant for the current investigation.

The most reliable ERP component elicited by syntactic garden path sentences is the P600, a posterior positivity between 600 and 1000 msec after onset of the word causing the anomaly (Osterhout & Holcomb, 1992). This component has also been elicited by various kinds of syntactic violations (Friederici, 2002) and has been interpreted as a correlate of additional processing costs during structural reanalyses (Friederici, 2002) or syntactic integration (Kaan & Swaab, 2003).

Unlike the P600, the N400 component is thought to primarily reflect difficulties during the processing of lexically bound semantic information. The N400 is a slightly right-lateralized negativity with centro-parietal scalp distribution that peaks around 400 msec. Initially, it was found to be elicited by implausible words such as the final word of the sentence "He spread the warm bread with socks" (Kutas & Hillyard, 1980). Subsequent work suggested that the N400 is a default response to content words which is modulated by the respective context (Kutas, van Petten, & Kluender, 2006). In other words, an increased N400 amplitude is an index of processing difficulty while retrieving and integrating a word in terms of its conceptual meaning. Whereas implausible theta roles often elicit N400s, a combined N400–P600 pattern is found if an NP does not receive a theta role (e.g., agent, patient) at all, for example, an object NP appearing after an obligatorily intransitive verb (Friederici & Frisch, 2000). As will become clear, both the N400 and the P600 are relevant to the present study because garden-path effects may interfere with lexical–semantic as well as syntactic integration.

More to the point, a number of ERP studies have shown that prosodic information is used in parsing as soon as it is available to the listener, whether in conditions containing a mismatch between prosodic and syntactic information (Mietz, Toepel, Ischebeck, & Alter, 2008; Eckstein & Friederici, 2006; Steinhauer et al., 1999) or in conditions containing syntactic ambiguity (Kerkhofs, Vonk, Schriefers, & Chwilla, 2007). Of particular relevance to the present investigation is a study by Steinhauer et al. (1999), which

demonstrated the guiding role of prosody in parsing by eliciting a prosody-induced garden path effect reflected in a biphasic N400–P600 ERP pattern. Their study included three conditions, the first two of which were well-formed German sentences containing a temporary attachment ambiguity. As in Kjelgaard and Speer's (1999) behavioral experiment, the critical third condition created a mismatch between the prosodic and the syntactic structure to probe prosody effects on parsing. Through digital cross-splicing of the speech signals, mismatching prosody and syntax were combined, resulting in a local violation of the verb's intransitive argument structure. In line with other ERP work, the N400 was taken to reflect lexical processing related to the verb argument structure violation in the stimuli, whereas the P600 reflected the subsequent structural reanalysis (Friederici & Frisch, 2000). The study provided the first electrophysiological evidence that prosodic boundaries can override the LC principle and change the initial parsing preference toward an EC analysis. Regrettably, the authors did not include a corresponding EC garden path condition that could elucidate the BDH. However, a more recent study by Bögels, Schriefers, Vonk, Chwilla, and Kerkhofs (2010), investigating similar sentences in spoken Dutch, did include EC conditions that lacked a boundary (along with LC conditions containing a superfluous prosodic break). Although their ERP analyses focused primarily on differences between object- and subject-control verbs (that are irrelevant in the present context), their data can, nevertheless, shed light on the BDH. Most importantly, whereas they reported systematic ERP garden path effects (N400s) for superfluous boundaries in LC structures (similar to Steinhauer et al., 1999), no such effects were found following missing boundaries in EC garden path sentences. The authors interpreted this pattern as evidence for a default preference for EC rather than LC, contrary to the observations of Steinhauer et al. (and thus, implicitly, contrary to the core assumptions of the garden path model). However, as these findings are perfectly in line with Steinhauer and Friederici's (2001) comma experiment, they may rather be viewed as a first indication that the BDH is, indeed, applicable to spoken language as well (see also Discussion).

The CPS Component

In the same study by Steinhauer et al. (1999), a distinct ERP component marking the perception of a prosodic boundary was identified and labeled the closure positive shift (CPS). This component was consistently recorded at the closure of prosodic phrases, and was characterized by a large positive-going waveform distributed bilaterally with a duration of about 500 msec (Steinhauer, 2003; Steinhauer et al., 1999). As the CPS profile in ERPs may resemble other ERP components (such as P600s or word onset components), Steinhauer (2003) described a number of criteria as to how these can be distinguished from a CPS; the interested reader is referred to that paper for further details. The CPS has since been replicated in German, Dutch, Japanese,

Chinese, and Korean (Bögels et al., 2010; Li & Yang., 2009; Wolff, Schlesewsky, Hirotsani, & Bornkessel-Schlesewsky, 2008; Kerkhofs et al., 2007; Pannekamp, Toepel, Alter, Hahne, & Friederici, 2005; Hwang & Steinhauer, in press), and has been argued to exclusively reflect the perception of prosodic boundaries, as it was elicited by speech signals that maintained prosodic information but were lacking both semantic and syntactic content (Pannekamp et al., 2005), and even by hummed sentences, lacking all types of segmental content (Pannekamp et al., 2005; Steinhauer & Friederici, 2001). This seems consistent with Beckman's (1996) proposal that prosody is, in fact, "a complex grammatical structure that must be parsed in its own right" (see also Fodor, 1998).

The Current Study

The first goal of our study was to replicate and extend the findings of Steinhauer et al. (1999) using English stimuli. To date, neither CPS data from English nor ERP data for the classical EC/LC garden path are available. Through digital cross-splicing of lexically identical but prosodically different initial sentence fragments, we were able to create two critical conditions that would presumably generate garden path effects based exclusively on their prosodic structure. Due to the nature of English EC and LC constructions, an interesting prosodic manipulation is possible. That is, in addition to examining the effect of a superfluous prosodic boundary as in Steinhauer et al.'s German stimuli, we can test the effect of the inappropriate absence of a prosodic boundary, and thereby the predictions of the BDH. Thus, unlike previous studies on conflicting prosody that introduced EC prosody into LC sentences and vice versa, we created an EC version without any boundaries (Condition C in Table 1) and an LC version that contained two boundaries, the first of which was superfluous (Condition D).

The EC sentence without boundaries can be viewed as an auditory analogue of the classical EC garden path in reading studies [see (1b) above]. Although garden path effects were predicted in both conditions, the respective on-line ERP patterns were expected to also reflect (a) the exact position of processing difficulties and (b) the qualitative and quantitative differences between the two types of garden paths. An overall advantage of LC over EC (Kjelgaard & Speer, 1999) could result in an attenuated P600 amplitude or duration for the LC sentences in (D). Alternatively, the BDH (Steinhauer & Friederici, 2001) would predict a smaller P600 for the EC garden path in (C), as only (D) requires the deletion of an existing prosodic boundary.

METHODS

Participants

Twenty-six undergraduate students from McGill University (13 women, age range = 18–25 years) were recruited by advertisement and paid for their participation. All were right-handed (Edinburgh Handedness Inventory; Oldfield,

Table 1. Sample Stimuli of the Four Experimental Conditions Marked with Cue Points

Condition	Sentence					↓ Lexical Disambiguation ↓		
	Conj	NP1	VP1	P1	NP2	P2	NP3	VP2
	1	2	3	4.....5	6	7	8	9
A (LC)	When	a bear	is approaching	the	people	#	the dogs	come running
	1	2	3	4	5	6	7.....9	
B (EC)	<i>When</i>	<i>a bear</i>	<i>is approaching</i>	#	<i>the people</i>			<i>come running</i>
C* (A1–B2)	When	a bear	is approaching	the	people			come running
D* (B1–A2)	<i>When</i>	<i>a bear</i>	<i>is approaching</i>	#	<i>the people</i>	#	the dogs	come running

Conditions C and D were derived from A and B by cross-splicing and inherited the corresponding cue points. See running text for details. The vertical line at cue point # 6 marks the splicing point. As indicated by arrows (↓), lexical disambiguation of the sentences was possible on the words following NP2 (i.e., NP3 “the dogs” in A and D and VP2 “come running” in B and C); these regions are relevant for the garden path effects in C and D.

1. Conditions A (in regular font) and B (in italics) are well-formed sentences. Conditions C and D are prosody–syntax mismatch conditions derived by means of cross-splicing; the asterisk (*) indicates the linguistic anomaly.

2. # marks a prosodic boundary.

3. The dotted vertical line represents the splicing point.

4. The lexically disambiguating elements correspond to the first words following the ambiguous NP2 (i.e., *the dogs* in A and D, and *come running* in B and C).

1971) native speakers of English with no known history of hearing impairment or brain injury. Prior to their participation, each subject gave written informed consent. Six subjects (2 women) were later excluded from further analysis due to EOG and movement artifacts exceeding 40% of the trials in one or more of the four main conditions.

Materials

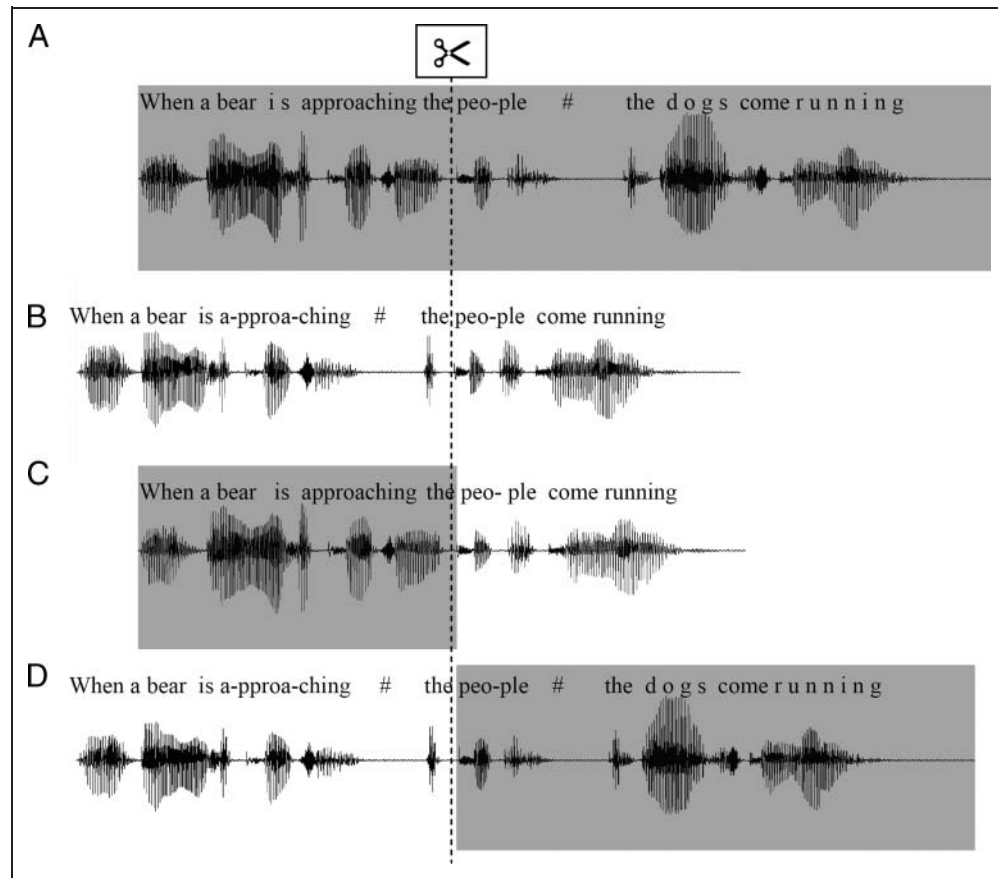
Four experimental conditions were created in two steps. First, 40 EC and LC sentence pairs in English (A and B in Table 1) were either adopted from previous studies (24; Walker et al., 2001; Kjelgaard & Speer, 1999), or were constructed anew following the same schema (16). The first VP in each sentence pair contained an optionally transitive main verb (e.g., *is approaching*), compatible with both LC and EC structures. As corpus analyses of the 40 verbs revealed an overall bias toward transitive use, and because we were interested in structural parsing preferences, we used the verbs in progressive aspect (“is approaching”) rather than present or past tense (“approaches”/“approached”). Progressive verb forms have previously been shown to reduce such a transitivity bias without abolishing the structural garden-path effects (Frazier, Carlson, & Clifton, 2006). The 80 sentences in Conditions A (LC) and B (EC) were recorded by a male English native speaker in a sound-attenuating booth (44.1 kHz sampling rate, 16-bit amplitude resolution [Marantz digital recorder PMD670]). Both conditions were produced with normal (cooperating) prosody, that is, with a boundary after the first verb in B, and a boundary after NP2 (“the people”) in A (see Table 1 and Figure 2).

In a second step, the two garden path conditions with conflicting prosody were derived from A and B by means of digital cross-splicing such that Condition C consisted of

the initial portion of Condition A and the final portion of Condition B, and vice versa for Condition D (Table 1 and Figure 2). Using cross-spliced versions rather than independent recordings of C and D has the major advantage that the resulting Conditions C and D contain the exact same physical speech signals as their control Conditions A and B. The splicing point was selected at the beginning of NP2 (e.g., “the people”), and thus, between the two boundary positions, such that Condition D inherited both boundaries, whereas Condition C contained no prosodic boundary. As aforementioned, Condition C can therefore be viewed as equivalent to the classical EC garden path condition in reading studies (see Figure 1B above). Its revision would require the establishment of a new boundary but not the deletion of an already existing boundary. Conversely, revision of D would require the deletion of a boundary but not the creation of a new one. If Kjelgaard and Speer’s (1999) hypothesized on-line processing advantage for LC holds, one might expect a stronger garden path effect for Condition C. Alternatively, if the necessity of deleting prosodic boundaries is a major factor contributing to processing difficulties, then Condition D would be expected to be much harder than C. In order to ensure that no audible amplitude shifts would be detected by the listeners, each sentence pair was spliced at the fricative “th” of the determiner preceding noun2 (e.g., *the people*).⁴ In total, 160 experimental stimuli were generated (in 4 conditions of 40 sentences each).

To allow a precise time-locking between relevant events in the speech signals (e.g., word or pause onsets) and the recorded EEG, nine cue points were inserted into each speech file of Conditions A and B, marking the constituents and the splicing point in these sentences (see Table 1). These cue points were inherited by Sentences C and D during the

Figure 2. Waveforms of sample sentences in all four conditions A–D. The vertical line below the scissors symbol indicates the splicing point. Speech signals on gray-shaded background were derived from Condition A, whereas those on white background correspond to Condition B. The hash mark (#) indicates prosodic boundary positions.



splicing. The cue points were essentially recorded along with the EEG and also used to determine and compare the duration of constituents in A and B as part of the acoustic–prosodic speech analysis (see below). In addition to the 160 experimental sentences, 160 unrelated filler sentences (in 4 conditions of 40 sentences each) were generated to prevent subjects from developing a strategy to recognize the critical mismatch conditions. Two of the filler conditions included well-formed stimuli (e.g., *The man hoped to enjoy the meal with friends*) and two filler conditions contained phrase-structure violations (e.g., *The man hoped to *meal the enjoy with friends*).

The 160 experimental and 160 filler sentences were evenly distributed across four blocks of 80 trials and were intermixed within each block in a pseudorandomized order. The following constraints were met: (1) each block contained 10 sentences in each of the four experimental conditions A, B, C, and D as well as the four filler conditions; (2) each block contained exactly one version (condition) of each sentence; (3) there was no consecutive repetition of the same condition; (4) match and mismatch conditions (in both experimental and filler sentences) were evenly distributed across each block; (5) no more than two experimental or two filler sentences occurred in a row, and no more than three correct or three anomalous sentences occurred in a row; (6) to minimize strategic processing effects, pseudorandomization within blocks also prevented (a) consecutive presentation of semantically related sen-

tences, (b) repetition of sentence initial conjunctions, and (c) clusters of particularly long or short sentences. Permutations of these four blocks were then used to form four experimental lists, using a Latin square design to counterbalance block order across lists. In order to further rule out any sequence effects, four additional “mirror” versions of each list were created by reversing both the block order and the sentence order within each block. Thus, a total of eight experimental lists were created and evenly assigned across male and female subjects.

Prosodic (Acoustic) Differences

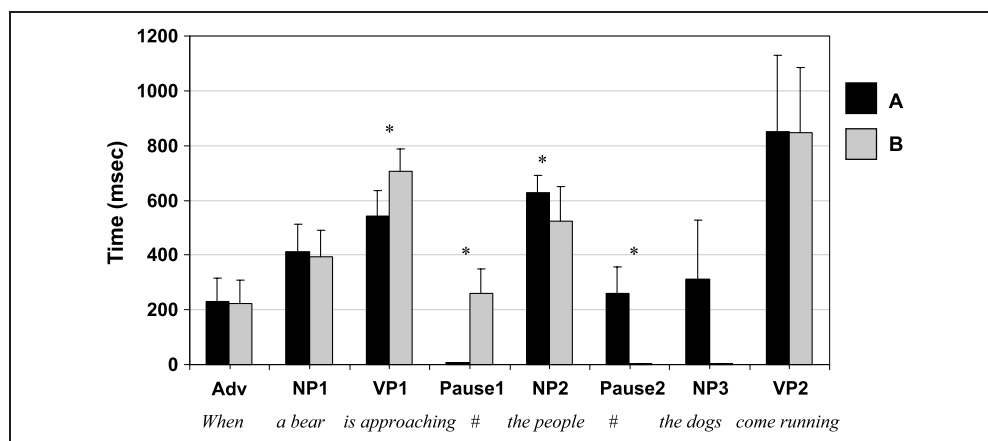
To ensure that the sentence conditions differed in terms of their prosodic structure as expected, we extracted (i) word/phrase and pause duration; (ii) pitch minima, maxima and averages; and (iii) amplitude/intensity information for each constituent marked by cue points (see Table 1) from the speech files of Conditions A and B and subjected them to statistical analyses.⁵ Similar to previous studies (e.g., Bögels et al., 2010; Kjølgaard & Speer, 1999; Steinhauer et al., 1999), we found the expected prosodic differences between Conditions A and B in duration measures (Figure 3). Contrasts of the constituent durations in A and B showed that the early boundary in B was marked by a highly significant preboundary lengthening of VP1 (*was approaching*) and a pause after VP1, whereas NP2 was lengthened and followed by a pause at the second boundary position in A

(all differences: $p < .0001$). The duration of the other constituents did not differ significantly between A and B. Statistically significant differences between A and B in terms of pitch and amplitude measures were not observed (all $ps > .05$), most likely due to (a) a strong reliance of our speaker on durational boundary markers and (b) inconsistent types of boundary tones across sentences. Kjelgaard and Speer's (1999) finding of such effects may have to do with the fact that their speaker was instructed to produce specific intonation contours, whereas our speaker was not. Note that Bögels et al. (2010) also reported significant prosodic condition effects for duration measures only.

Procedure

Participants sat in a comfortable chair approximately 80 cm in front of a computer monitor inside an electromagnetically shielded, sound-attenuating chamber (IAC America Inc., New York, NY), and listened to spoken sentences presented binaurally via insert earphones (Etymotic Research, Elk Grove Village, IL). Subjects were asked to press the left button of a computer mouse if the sentence sounded natural/acceptable or the right button if the sentence sounded unnatural/unacceptable (acceptability judgment task). Subjects were not given explicit examples of acceptable/unacceptable sentences. Each trial began when a fixation cross appeared on the screen 1000 msec before sentence presentation, which remained visible until the end of the sentence. Subjects were instructed to fixate the cross and avoid any blinking, eye movements, and body movements during sentence presentation. Then, a response prompt (i.e., "Natural or not?") appeared on the screen until a mouse button was pressed or the maximum response time (5 sec) had elapsed (whichever came first). Next, a prompt ("!!!") appeared on the screen for 1500 msec, indicating the interval during which the subjects were instructed to blink their eyes before the next trial began. Each session began with a short practice block containing 10 sentences, after which further questions were clarified if necessary.

Figure 3. Duration measures for constituents and pauses confirm the predicted acoustic-prosodic differences between Conditions A (black bars) and B (gray bars). Whereas VP1 and Pause 1 were significantly longer in "early closure" Condition B (marking the first boundary), NP2 and Pause2 were longer in "late closure" Condition A (marking the second boundary). Note: NP3 occurred only in Condition A. Asterisks (*) indicate significant differences.



EEG Recording

EEG was continuously recorded (500 Hz/32 bit sampling rate; Neuroscan Synamps2 amplifier) from 19 cap-mounted Ag/AgCl electrodes (Electro-Cap International, Eaton, OH) placed according to the standard International 10–20 System at the following sites: Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, O2. Vertical and horizontal eye movements (EOG) were recorded from bipolar electrode arrays placed above and below the left eye and at the outer canthus of each eye, respectively. All EEG electrodes were referenced against the right mastoid; and an electrode located halfway between Fpz and Fz served as the ground. Electrode impedances were kept below 5 kΩ.

Behavioral Data Analysis

Acceptability ratings were computed as the percentage of accepted sentences separately for each condition. Very early responses (<200 msec) and very late responses (>5000 msec) were excluded from the analyses. Response times were computed separately for accepted and rejected trials in each of the four main conditions. Data were subjected to repeated measures ANOVAs with the factors prosody (2) and violation (2).

ERP Data Analysis

The EEG data were analyzed using EEProbe software (ANT, The Netherlands). Single-subject averages were computed separately for each experimental condition following a pre-processing analysis which included filtering (0.16–30 Hz bandpass), artifact rejection and detrending. All EEG epochs contaminated with EOG and movement artifacts exceeding a 30-μV threshold were excluded from the averaging procedure. Only the data of those subjects with a minimum of 25 trials per condition entered the statistical analysis. On average, 19% of the trials were lost due to artifacts; this percentage did not differ across conditions. In order to compensate for the latency variability of speech signals, averages time-locked to various events were computed (based on

the cue points shown in Table 1); moreover, testing the robustness of effects occasionally required the use of multiple baseline intervals. Details are described in the Results section. ERP components were quantified by means of amplitude averages in representative time windows (see Results section). Global analyses of variance (ANOVAs) for repeated measures were carried out for the data in each time window, separately for 3 midline and 12 lateral electrode sites. The ANOVAs for ERP data from the lateral electrodes included two condition factors, namely, prosody (sentence fragment containing an IPh boundary vs. sentence fragment without an IPh boundary; see also Table 2)⁶ and violation (prosody–syntax match vs. prosody–syntax mismatch), as well as three topographical factors, namely, hemisphere (right vs. left), laterality (medial vs. lateral⁷), and anterior–posterior (frontal vs. central vs. posterior). ANOVAs for the midline electrodes included the same factors except for the topographical factors hemisphere and laterality. Additional ANOVAs followed up on significant interactions. In order to address violations of sphericity, the Greenhouse–Geisser correction was applied to all repeated measures with more than one degree of freedom in the numerator, for which we report original degrees of freedom and corrected *p* values.

RESULTS

Behavioral Data

Acceptability rates and response times of the behavioral task are displayed in Figure 4 and Table 3, respectively. Whereas naturally spoken LC and EC sentences in Conditions A and B were equally well accepted—87.5% and 87.2% of the time, respectively ($F < 1$), and did not differ in response times either ($F < 1$)—the classical garden path Condition C was accepted in only 53.3% of the trials, and Condition D, with its two boundaries, was rated as even less acceptable (28.0%). The difference in acceptability between C and D, and all other pairwise comparisons with A and B, were highly significant [all $F(1, 19)$ values > 22 , $ps < .0001$]. Moreover, subjects were faster in accepting A and B than the cross-spliced Conditions C and D [$F(1, 14^8) = 24.86$, $p < .0001$], but tended to need more time to reject

Table 2. Matched Condition Pairs before and after the Splicing Point, Illustrating Which Conditions Shared Identical Speech Signals and Were, therefore, Contrasted in the ERP Analyses

	Before the Splicing Point	After the Splicing Point
Boundary present	B (EC)/*D(LC)	A (LC)/*D(LC)
Boundary absent	A (LC)/*C(EC)	B (EC)/*C(EC)

Early closure (EC) and late closure (LC) indicate which syntactic analysis was required based on the lexical information.

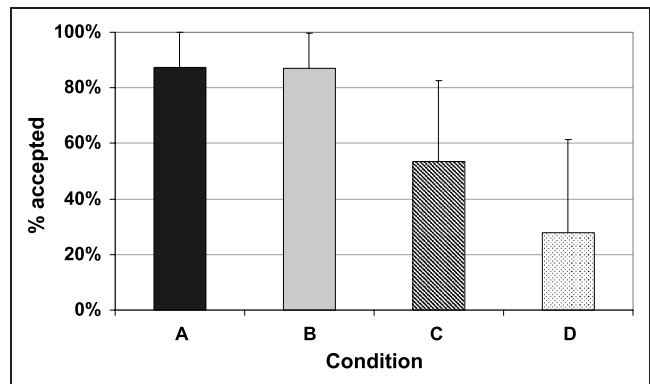


Figure 4. Acceptability ratings (% accepted trials) per condition.

Whereas conditions with cooperating prosody (A and B) were accepted 87% of the time, the prosody-induced garden path in Conditions C and D resulted in processing difficulties. Acceptability was particularly low in Condition D, which required the listeners to mentally delete a boundary.

stimuli in Condition C compared to Condition D [$F(1, 19) = 3.62$, $p < .08$]. Only in Condition D did participants need more time to accept sentences than to reject them, although this numerical difference did not reach significance ($p > .05$).

ERP Data

Because the ERPs time-locked to the sentence onset do not provide the desired precision due to variability in word and constituent length across our 40 sentences and, as a result of cross-splicing, the relevant control condition comparisons differ before and after the splicing point, we will address CPS and garden path effects time-locked to local events and compared to the appropriate matched controls. As boundary effects will help us understand the garden path effects, we will first present the CPS findings.

ERP Effects at Boundary Positions

Figure 5A and B illustrates the ERPs time-locked to the offset of words immediately preceding the two pauses. In

Table 3. Response Times (msec) for Accepted and Rejected Trials Per Condition

Condition	Accepted Trials		Rejected Trials	
	Mean	SD	Mean	SD
A	683	334	1180	736
B	702	324	998	632
C	929	440	999	529
D	944	422	755	444

Figure 5. (A) CPS1 in Conditions B and D at the first boundary. Grand-average ERPs of all four conditions are time-locked to the offset of the first verb (“is approaching”; vertical lines at 0 msec), using a baseline of -500 to 0 msec. Conditions B (gray, dotted) and D (black, dotted) evoke a closure positive shift (CPS) at pause onset, which is directly preceded by a broadly distributed negativity. Conditions A (gray, solid) and C (black, solid), which do not contain a boundary at this position, elicit a slow frontal negativity between 200 and 700 msec.

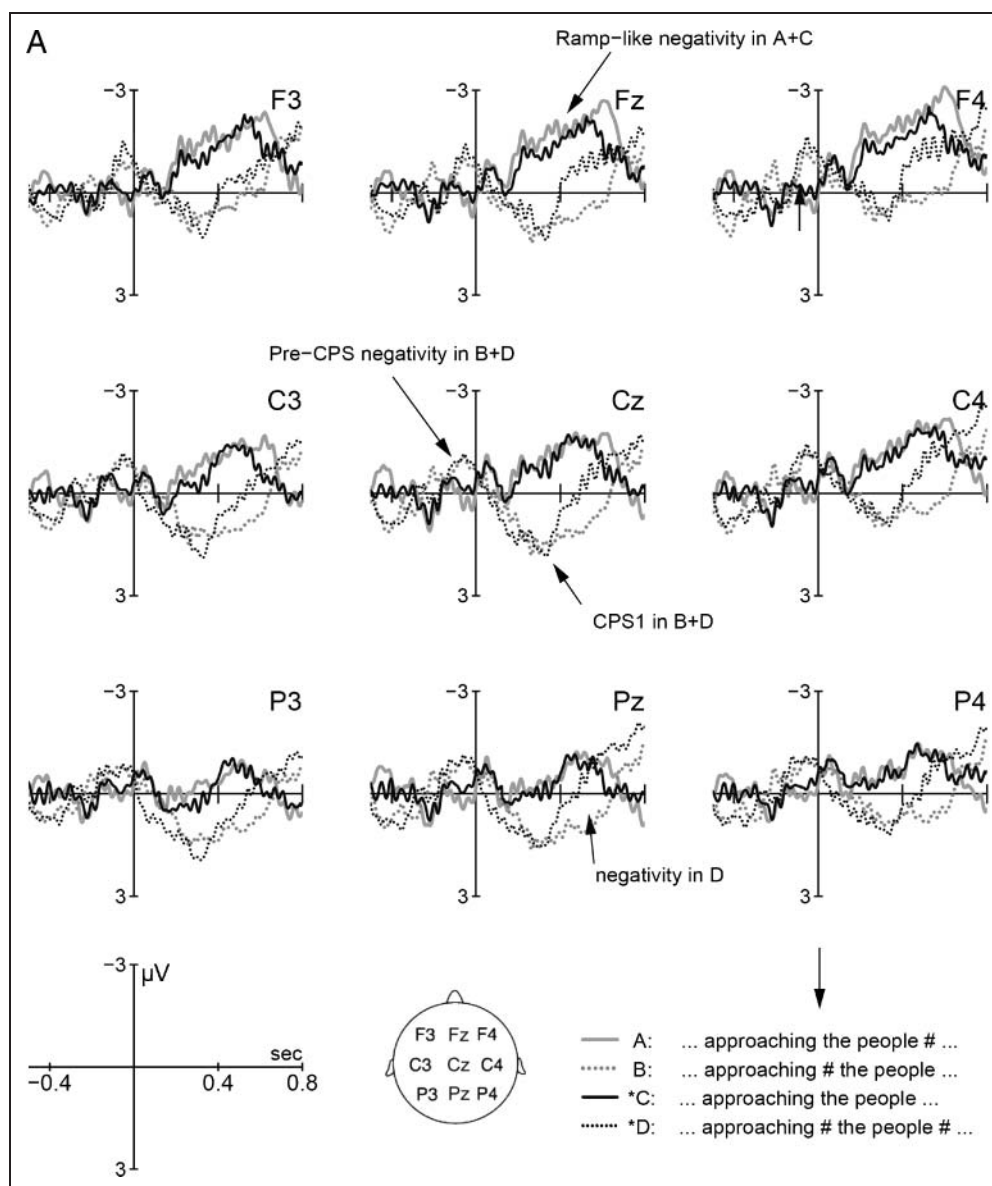
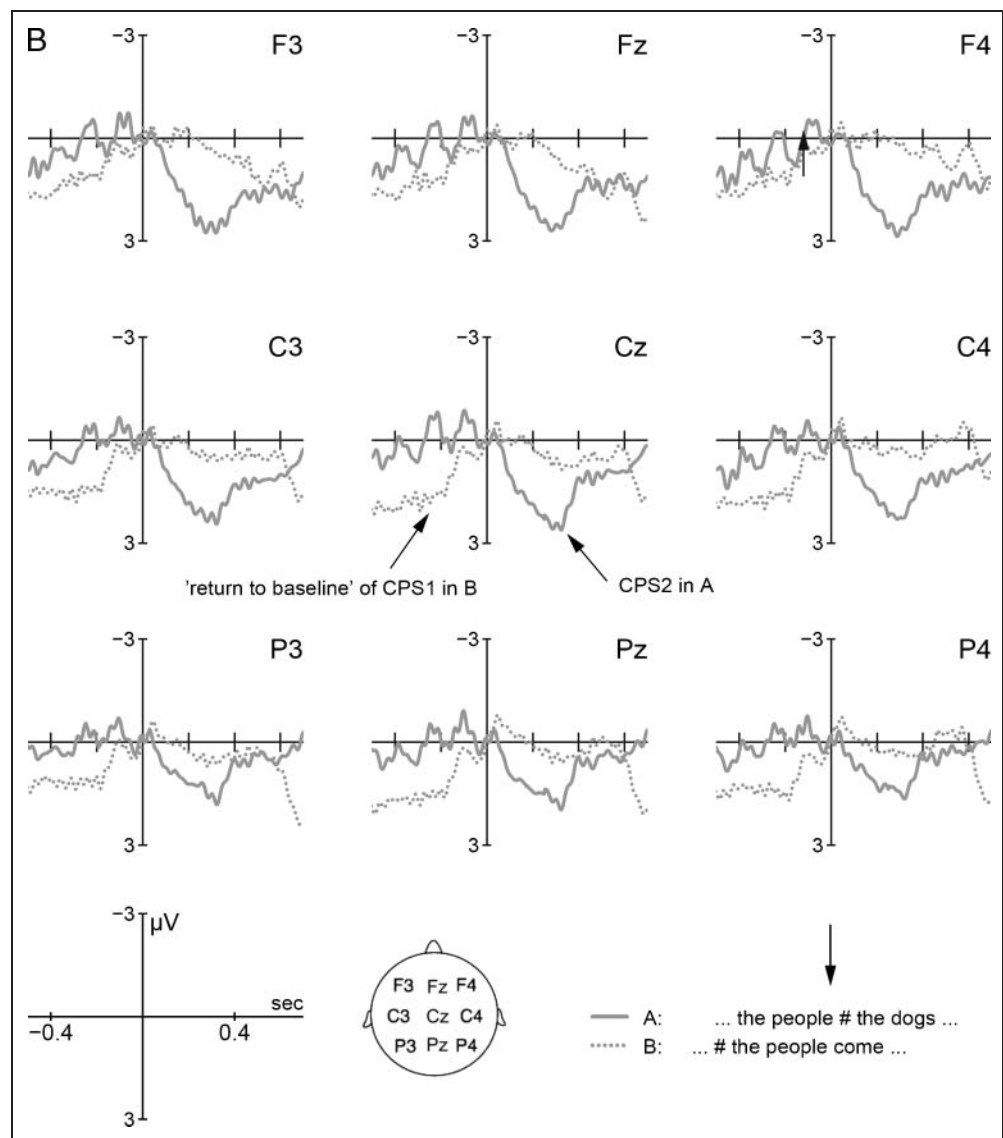


Figure 5A, ERPs in all four conditions are shown for the first boundary position at verb1 offset (“is approaching_”). As this position lies prior to the splicing point, matched sentence pairs A/C and B/D show virtually the same pattern. Most importantly, we can see that the two conditions containing a prosodic boundary (B and D) both show the expected closure positive shift immediately starting at the offset of the verb. This CPS1 has a broad distribution with a central amplitude maximum, peaks at about 300 msec and returns to baseline between 500 and 700 msec. This latter observation is important, as previous studies describing CPS components (e.g., Kerkhofs et al., 2007; Pannekamp et al., 2005; Steinhauer et al., 1999) did not explicitly address its return to baseline. This negative shift continues even after crossing the baseline and also includes the N400-like effect in D (see below). Unlike B and D, Conditions A and C display a fronto-central negative shift between 200 and 700 msec, peaking around 550 msec. This peak is larger in

A than C, likely due to differences after the splicing point. A last relevant observation concerns the negativity directly preceding the CPS in B and D which reaches its peak amplitude at about -50 msec. If we adopt the notion that this peak marks the actual onset of the CPS in B and D, the CPS must indeed have been triggered by events that occurred prior to the pause onset.

All of the above observations were confirmed by statistical analyses. Due to the variability of speech and music signals, baselines in auditory ERP experiments tend to be less robust than in reading studies, such that multiple analyses using different baselines, or baseline-independent measures, are advantageous to demonstrate robust effects (e.g., Knoesche et al., 2005; Steinhauer, 2003). We thus conducted analyses using three different baseline intervals: (a) -500 to -150 msec, (b) -500 to 0 msec, and (c) -50 to 50 msec. All of them revealed consistent effects for both the CPS and the preceding negativity. Here we report analyses

Figure 5. (B) CPS2 in Condition A (compared to B) at the second boundary. Grand-average ERPs are time-locked to the offset of NP2 (“the people”; vertical lines at 0 msec), using a baseline of –50 to +50 msec. Condition A (gray, solid) evokes a CPS at pause onset, which returns to baseline after some 500 msec. Condition B (gray, dotted), which does not contain a boundary at this position, does not show this CPS. Instead, at about –200 msec, we see the return-to-baseline of its CPS1 at the first boundary position (see also Figure 5A).



with the baseline from –500 to 0 msec that also underlies the plots in Figure 5A.

The Negativity Preceding CPS1

Between –200 and 0 msec, we found a main effect of Prosody at both lateral [$F(1, 19) = 15.95, p < .0001$] and midline electrodes [$F(1, 19) = 25.60, p < .0001$]. At lateral electrodes, additional interactions of Prosody with factors Laterality [$F(1, 19) = 13.00, p < .002$] and Laterality \times Hemisphere [$F(1, 19) = 4.46, p < .05$] indicated that the negativity was most prominent at medial electrodes and larger over the right hemisphere. A similar small right-lateralized negativity preceding the CPS was also observed by Bögels et al. (2010), who referred to it as a “reversed effect.”

The CPS1 in B and D

Analyses at lateral sites between 0 and 600 msec revealed a main effect of Prosody [$F(1, 19) = 6.16, p < .03$] and a highly

significant Prosody \times Laterality interaction [$F(1, 19) = 39.47, p < .0001$], indicating that the difference between Conditions A/C and B/D was larger over medial [$F(1, 19) = 10.90, p < .004$] than lateral/peripheral ($p > .17$) electrodes. At midline electrodes, we found both a main effect of Prosody [$F(1, 19) = 11.88, p < .003$] and a marginal Prosody \times AntPost interaction [$F(2, 38) = 3.81, p < .06$], pointing to a maximum at Cz. More fine-grained analyses of consecutive 100-msec time windows showed that the prosody-related differences were most prominent during the first 500 msec. The absence of any effects involving the factor violation ($F_s < 1$) underlines the fact that the matched condition pairs (A/C and B/D, respectively) elicited indistinguishable ERPs during the first 500 msec. Between 500 and 700 msec, an additional Prosody \times Violation \times Hemisphere interaction [$F(1, 19) = 4.43, p < .05$] reflected the somewhat right-lateralized N400-like negativity in Condition D.⁹ As the ramp-like frontal negative shift in A and C may have contributed to the differences between B/D and A/C, we also conducted single-sample *t* tests to determine when and at

which electrode sites each of the four conditions differed significantly from the baseline. These t tests demonstrated that the CPS in B and D was highly significant at central and posterior electrodes during the first 200–500 msec [Cz, Pz, C3, P3, P4, T5; all $t(19) > 2.5, p < .02$].¹⁰ In contrast, the negative shift in A and C reached significance only after 400 msec, was constrained to frontal and central electrodes, particularly over the right hemisphere, and reached its highest significance in Condition A at electrodes F4 and F8 between 500 and 700 msec [$t(19)$ values $> 4.75, p < .0001$]. Combined, these analyses indicate that both the CPS at the first boundary position in Conditions B/D and the frontal negativity in A/C contributed to the prosody effects, although with distinct latencies and scalp distributions.

The CPS2 in Condition A

Both the ANOVA comparing Conditions A and B and the single-sample t tests for Condition A confirmed the significance of the CPS in Condition A (Figure 5B). The analysis between 0 and 600 msec revealed a main effect of Prosody at both lateral [$F(1, 19) = 4.95, p < .04$] and midline electrodes [$F(1, 19) = 4.76, p < .05$] as well as a Prosody \times Laterality interaction at lateral sites [$F(1, 19) = 5.58, p < .03$], again indicating larger amplitudes near the midline. As with the first CPS, the effect was most reliable between 100 and 400 msec [e.g., Prosody \times Laterality: $F(1, 19) = 14.65, p < .002$]. t tests for the same time interval demonstrated that the CPS in Condition A differed significantly from baseline at 13 out of 15 electrodes [with $t(19)$ -values ranging between 2.65 ($p < .03$) and 4.81 ($p < .0001$)]. In terms of both magnitude and reliability, the CPS was largest at frontal electrodes, particularly over the right hemisphere. As this pattern mirrors the distribution of the negative shift in A and C prior to this boundary (see Figure 5A), this negativity may have influenced the subsequent positive shift (see below). The small positive deflection in B did not reach significance ($p > .1$).

Comparison of CPS1 and CPS2

To examine the variability of CPS components, we directly compared the three CPS components in Conditions B, D, and A (factor condition) between 0 and 600 msec (as well as in smaller time windows) using a -50 to $+50$ msec baseline. Apart from shared effects of laterality [$F(1, 19) = 7.36, p < .02$], we found a significant Condition \times AntPost interaction [$F(4, 76) = 4.16, p < .005$]. Pairwise comparisons revealed that the CPS in A was consistently more frontal than those in Conditions B [Condition \times AntPost: $F(1, 19) = 5.27, p < .01$] and D [$F(1, 19) = 5.79, p < .007$], which did not differ from each other ($F < 1$). However, as shown in Figure 6, which contrasts the late CPS2 in A with the early CPS1 (averaged across Conditions B and D) using a baseline of -650 to -550 msec, these differences were primarily due to the frontal negative shift preceding the CPS in Condition A; selecting the baseline prior to onset of this nega-

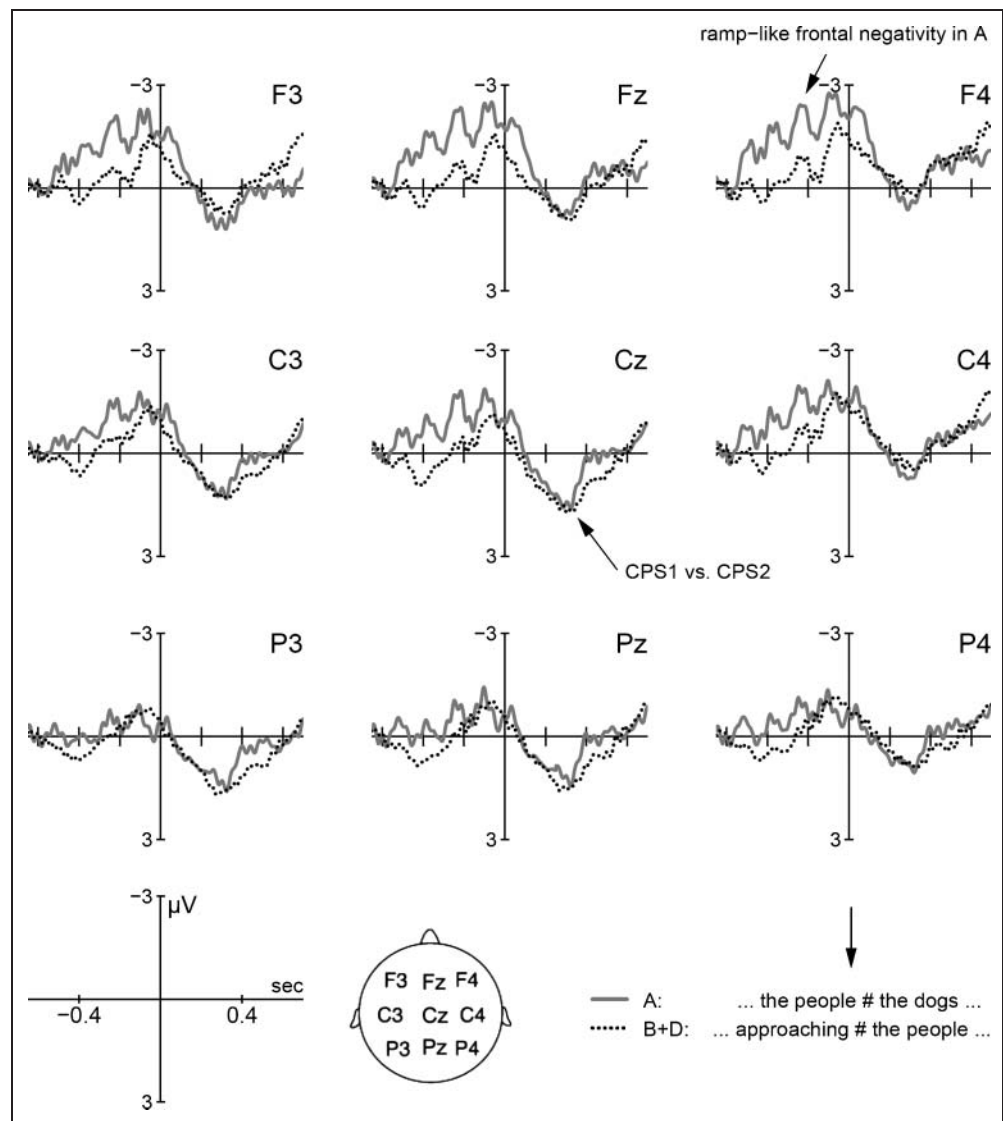
tivity virtually eliminates all distributional differences among the CPS components. In other words, at frontal electrodes, the positive shift in Condition A can be best described as a combination of both (a) the reset of the frontal negativity and (b) the CPS proper.

Garden Path Effects

Recall that after the splicing point, Condition C now shares the same speech signal with Condition B, whereas ERPs in Condition D were elicited by the exact same speech signal as those in Condition A (see Table 2). Figure 7 gives an overview of the ERPs in all four conditions from 500 msec before to 1800 msec after the splicing point. It provides information about the general profile and scalp distribution of the ERP components. The other three figures (Figures 8A–B and 9A) show selected pairwise comparisons among the four conditions at the midline and voltage maps for corresponding garden path differences between the two respective conditions (BD, AD, and BC). For these analyses, we chose a baseline interval of -500 to 0 msec in order to minimize the impact of the first CPS (in Conditions B and D) on the current comparisons. These CPS1 components can be seen between -200 and $+100$ msec relative to the splicing point, most prominently at central electrodes.

Between 200 and 700 msec, all four conditions display a broadly distributed negativity (see Figure 7). In B and D, it corresponds to the post-CPS1 negativity, whereas in A and C it reflects the negative shift preceding the second boundary position. Only in Condition D does the negativity reach amplitudes of more than $2 \mu\text{V}$ even at posterior electrodes. This is due to the additional N400-like effect in this condition, which is best illustrated in Figure 8A comparing Conditions B and D. Importantly, this negativity marks the first significant difference between Conditions B and D. Its onset latency of about 200 msec suggests that the ERP during the first 200 msec after the splicing point was still exclusively determined by the speech signal presented prior to it. After this point, the listener's brain had clearly begun to integrate the deviant prosodic information in D. With respect to data analysis, this finding means that after 200 or 300 msec, the control conditions for the cross-spliced conditions need to be swapped: The more valid control for Condition D should now be A, that for C should be Condition B.¹¹ At about 500 msec, the N400 in D abruptly changes into a large posterior positivity between 700 and 1400 msec, peaking at about 1000 msec. The second cross-spliced Condition C also elicits a positive-going waveform with posterior distribution, however, approximately 500 msec later than D. As we assume these positivities to reflect linguistic processing difficulties, we will henceforth refer to them as P600 components. The prosody–syntax mismatch (garden path) effects were analyzed in four time windows: Two windows were used to quantify the N400 in Condition D: from 150 to 450 msec to capture the early difference between D and B, and from 300 to 550 msec for

Figure 6. Direct comparison of CPS1 (collapsed across Conditions B and D, dotted line) and CPS2 in Condition A (solid line). Selection of an early baseline interval (−650 to −550 msec) reveals that the positive shift in A comprises both the “reset” of the ramp-like frontal negativity and the CPS proper. The actual CPS components at both boundary positions display strikingly similar morphology and distribution.



differences between D and A. For the P600 components in D and C, we selected nonoverlapping time intervals of 700–1300 and 1300–1700 msec, respectively.

The N400 in Condition D

The centro-parietal negativity in Condition D between 200 and 500 msec indicates that, at this point, the different speech signals in B and D started to affect the ERP. However, this does not necessarily imply a syntax–prosody mismatch (garden path) effect: Because the speech signal in D consists of the first part of B and the second part of A, and because both A and B also elicit negativities in this time range in absence of any mismatch, the N400-like effect may, in principle, be due to a superimposition of these other components. In order to justify an interpretation of the negativity in D in terms of a mismatch effect, we therefore have to demonstrate that it differs not only from the negativities of both A and B but also from the sum of these negativities. As summarized below, all of these compari-

sons did, in fact, confirm that the negativity in D cannot be explained by any combination of the negativities in A and B: It is not only larger in amplitude but also more posteriorly distributed than the other effects. Table 4 summarizes the analyses for the 150–450 msec time window, where significant interactions of Violation × Prosody × Laterality at lateral electrodes and of Violation × Prosody at the midline reflect the difference between D and B (and its absence in C vs. A; $p > .7$) in the global ANOVAs. The follow-up comparison of B and D showed the expected prosody effects and revealed that the difference was most prominent at Pz ($p < .004$), compatible with an N400. In contrast, the small difference between A and C at right anterior electrodes was reflected by a four-way interaction (Prosody × Laterality × AntPost × Hemisphere; $p < .03$). Additional analyses in a standard N400 interval of 300 to 550 msec replicated all of these differences and also demonstrated that Condition D differed at Pz from all three other conditions (A: $p < .03$; B: $p < .007$; C: $p < .003$). Last, to test the hypothesis of additive negativities underlying the N400, the difference wave

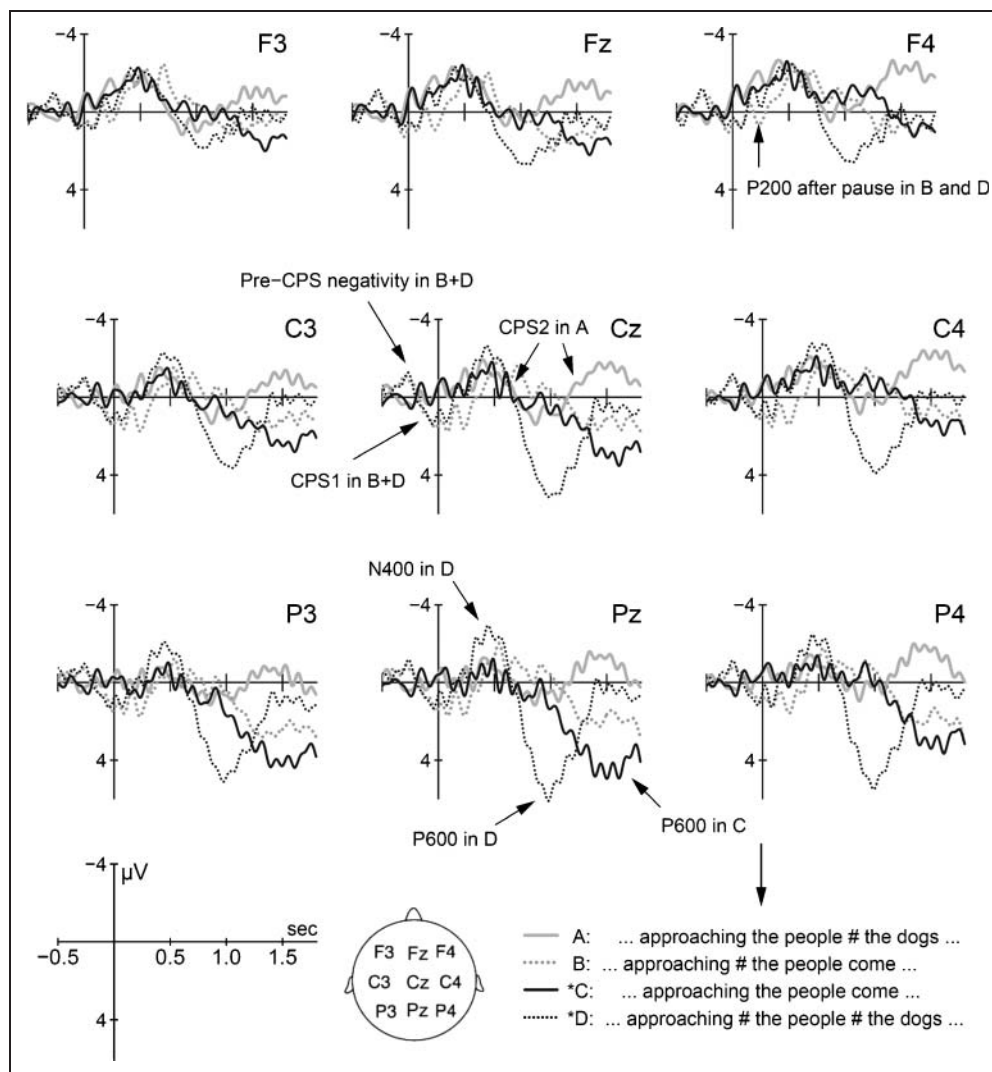
between D and B (i.e., the N400 effect) was contrasted to the negativity in A. Condition \times AntPost interactions at both lateral and midline electrodes reached significance in the 150–450 msec as well as the 300–550 msec intervals [all $F(2, 38) > 5, p < .02$], confirming the significantly more posterior distribution of the N400 in Condition D. Together, these analyses strongly suggest that this component reflects the first prosody–syntax mismatch effect.

The P600 in Condition D

The large parietal positive component following the N400 in Condition D differed significantly from all other conditions between 700 and 1300 msec and was somewhat larger over the right than the left hemisphere (see Table 5 for details). Compared to its matched control condition after the splicing point (i.e., Condition A), it continued to differ significantly until the end of the epoch (1800 msec; see Figure 8A and B). Two observations regarding its timing are worth mentioning. First, the P600 in Condition D has almost

the same onset, peak and offset latencies as the CPS2 in Condition A (Figure 8B). Therefore, it is impossible to determine whether or not the shared boundary at this position elicited a CPS in Condition D as well or not. However, it is also clear that, in addition to its much larger amplitude, the scalp distribution of the P600 was more posterior than that of the CPS in A [Violation \times AntPost: $F(2, 38) = 10.51, p < .003$], ruling out any common account for the two effects despite their similarity in timing. Secondly, the rather early occurrence of both the N400 and the P600 in Condition D suggests that these effects were elicited by the earliest available prosodic cues providing boundary information. Additional analyses time-locked to the offset of the pause, that is, the onset of NP3 (not shown), also confirmed that the P600 was elicited (and even reached half of its peak amplitude) before the end of the pause. That is, listeners encountered processing difficulties right at the second boundary (“When a bear is approaching #1 the people #2 ...”) and did not have to wait even for the third NP (“the dogs”). We will return to this issue in the Discussion section.

Figure 7. Grand-average ERPs of all four conditions time-locked to the splicing point (determiner “the” of NP2 *the people*), illustrating the prosody–syntax mismatch effects. Whereas Condition D elicits a biphasic pattern consisting of an N400 and a subsequent large P600-like positivity at the second boundary, Condition C evokes a P600 at the disambiguating second VP.



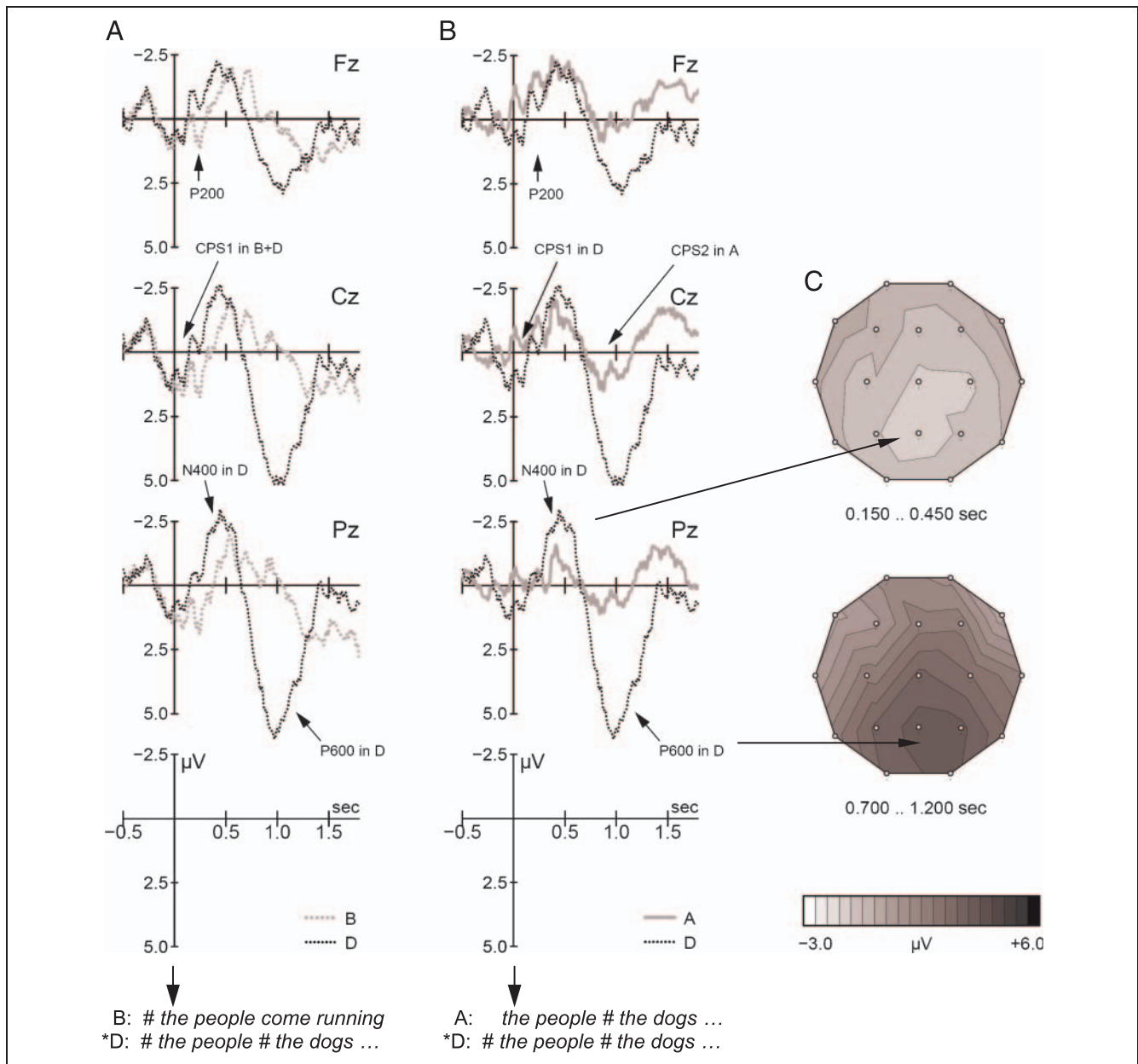


Figure 8. Garden path effects in Condition D at midline electrodes, time-locked to the splicing point (determiner “the” of NP2 *the people*). (A) The centro-parietal N400 in D reflects the first difference from Condition B at about 200 msec after the splicing point and likely reflects semantic-thematic processing difficulties. As NP2 (*the people*) is lengthened in D (but not in B), the earlier onset of the negative shift in D cannot be explained by word length differences. At Fz, onset N100–P200 components of the postboundary words in both B and D also illustrate that the much earlier CPS1 is completely independent of these components. (B) After the splicing point, ERPs in Conditions A and D were elicited by the exact same speech signals. The large P600-like positivity may reflect attempts to reanalyze the structure, which seemed to have failed in most trials. (C) Voltage maps of the difference waves illustrate the centro-parietal, slightly right-lateralized distribution of the N400 (D – B) and of the parietal P600 (D – A).

The P600 in Condition C

In contrast to Condition D, the P600 in Condition C was elicited rather late, that is, at the second VP (*When a bear is approaching the people come running*; see Figure 9A). This indicates that the listeners’ processing difficulties depended on the integration of the lexically disambiguating information of the verb, not just on the absence of the second pause. The P600 effect in Condition C was significant between 1400 and 1700 msec after the splicing point

and had a somewhat left-lateralized parietal distribution (Table 5). However, because Condition C and its matched control Condition B already differed earlier (i.e., between 500 and 1400 msec) due to the impact of the CPS1 in B and subsequent negativities, we also ran a comparison between ERPs in B and C time-locked to the onset of the second (disambiguating) VP, employing a baseline of –50 to +50 msec that eliminated earlier differences. This comparison, which largely resembles that in Figure 9A, is illustrated in Figure 9B. The corresponding analyses for

these averages were performed in a time window from 900 to 1100 msec and replicated all relevant effects. Significant interactions of violation, with the factors laterality and AntPost (p values $< .02$), underlined the medial and posterior distribution of the P600 [main effect of violation at Pz: $F(1, 19) = 6.48, p < .02$]. Thus, the P600 was found to be independent of earlier differences between Conditions B and C.

A last analysis concerned the question as to whether the P600 in Condition C varied in amplitude depending on whether the sentences were ultimately rated as acceptable or not.¹² This analysis, also computed for ERPs time-locked to VP2 onset, was not entirely conclusive. On the one hand, the P600 amplitudes of accepted trials (Pz: $4.44 \mu\text{V}$) and rejected trials (Pz: $5.04 \mu\text{V}$) in Condition C did not differ significantly from each other (e.g., at Pz: $F < 1$). On the other hand, only rejected trials differed significantly from control Condition B [Pz: $F(1, 19) = 4.73, p < .05$], whereas the dif-

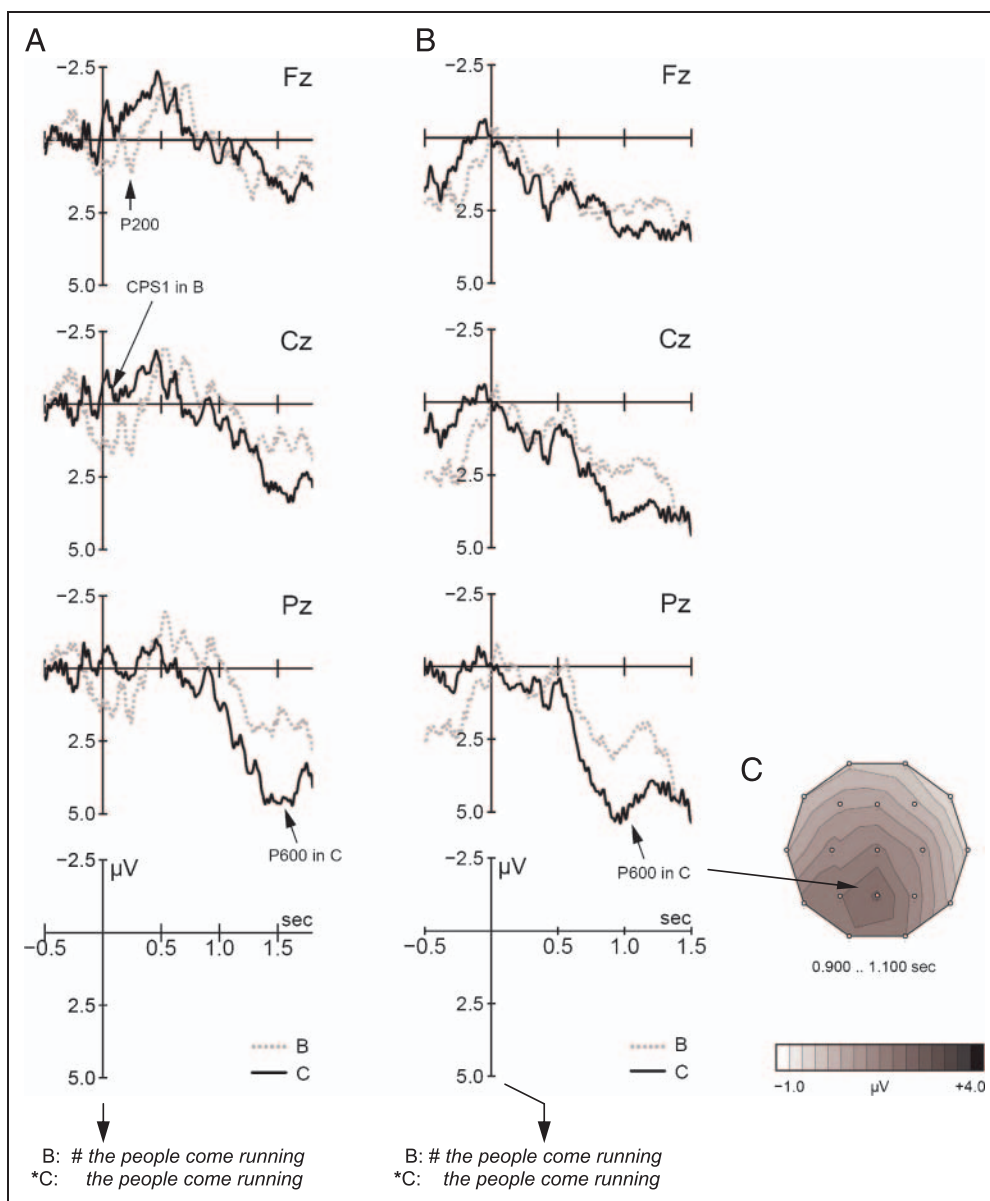
ference between B and accepted trials in C was only marginally significant [$F(1, 19) = 2.99, p < .1$]. This pattern as a whole was likely due to individual variability and the relatively small number of trials per subcondition. It suggests that the P600 effect was somewhat less reliable in accepted trials.

To summarize, Condition D elicited a biphasic N400/P600 pattern right at the second boundary, whereas Condition C displayed a P600 at the lexically disambiguating second VP, whose amplitude was numerically larger in rejected than accepted trials.

DISCUSSION

The present ERP study investigated the on-line processing of spoken English garden path sentences (early vs. late closure; EC, LC) with either cooperating or conflicting

Figure 9. P600 garden path effect in Condition C compared to matched control Condition B, time-locked to (A) the splicing point and (B) the onset of VP2 (*come running*). The plots in (B) are more reliable as they compensate for earlier differences between the conditions that were influenced by the CPS1 in B. (C) The voltage map of the P600 difference wave (C - B) shows a slightly left-lateralized parietal distribution.



prosody. The EC Condition C (with no audible boundaries) elicited a P600 on the disambiguating second VP, whereas the LC Condition D (with two boundaries) displayed an N400 followed by a large P600 at its second prosodic boundary. ERPs also reliably replicated the CPS at prosodic boundaries and shed new light on the profile and temporal fine structure of this component. We will now discuss these findings, in turn, starting with the garden path effects.

Garden Path Effects in Condition C

Compared to its matched control B, Condition C elicited a relatively small, but significant, parietal positivity on the

Table 4. N400 Effect in Condition D (150–450 msec Relative to Splicing Point)

Source	<i>df</i>	<i>F</i>	<i>p</i>
<i>1. Global ANOVA</i>			
Lateral electrodes			
Violation × Prosody	1, 19	4.10	.06
Violation × Lat	1, 19	7.28	.02
Violation × Antpost	2, 38	3.53	.06
Prosody × Hemi	1, 19	6.58	.02
Viol × Pros × Lat	1, 19	8.63	.01
Viol × Lat × Antpost	2, 38	4.99	.02
Pros × Lat × Antpost	2, 38	4.18	.03
Pros × Lat × Hemi	1, 19	4.64	.05
Pros × Lat × Antpost × Hemi	2, 38	6.46	.01
Viol × Pros × Lat × Antpost × Hemi	2, 38	3.37	.06
Midline electrodes			
Violation × Prosody	1, 19	7.10	.02
Prosody × Antpost	2, 38	5.56	.02
<i>2. Pairwise Comparison: D vs. B</i>			
Lateral electrodes			
Violation	1, 19	4.37	.06
Violation × Lat	1, 19	10.67	.005
Violation × Antpost	2, 38	5.13	.03
Viol × Lat × Antpost	2, 38	4.66	.02
Midline electrodes			
Violation	1, 19	8.57	.001
Fz	1, 19	4.78	.05
Cz	1, 19	8.58	.01
Pz	1, 19	10.92	.004

morphosyntactically (lexically) disambiguating second VP (VP2). We interpret this effect as a P600 component previously found to reflect syntactic processing difficulties in both garden path sentences (Osterhout, Holcomb, & Swinney, 1994) and syntactic violations (Friederici, 2002). The effect was predicted under the assumption that, without explicit prosodic boundary information, listeners would follow the LC principle (Frazier, 1987; Frazier & Fodor, 1978) and initially interpret the ambiguous NP2 as the object NP of the preceding verb.¹³ Only upon encountering the second verb did they realize that NP2 was actually needed as the subject of that verb, thus requiring a structural reanalysis (see Figure 1B). In contrast to the vast majority of previous studies exploring prosody-induced (EC/LC) garden path effects, Condition C did not include any misleading boundary information. Rather, the garden path was due to the lack of an early boundary, and thus, mirrored the classical EC garden path in reading studies (Steinhauer & Friederici, 2001; Frazier & Rayner, 1982). In spoken sentences, this reanalysis may involve the mental creation of a new boundary after verb1, which according to the boundary deletion hypothesis, should be rather easy. Our P600 finding extends previous behavioral data investigating similar structures (Walker et al., 2001; Kjelgaard & Speer, 1999) and supports the guiding role of the LC principle in spoken sentences that lack prosodic boundaries.

On the other hand, the rather small and local P600 effect also suggests that the structural revision was, indeed, relatively easy to carry out, thus confirming the prediction of the BDH. Ease of reanalysis corresponds to the relatively high acceptance of Condition C (compared to D). However, the difference in acceptance of 30% between C and its matched correct control Condition B also illustrates that listeners expect speakers to be unambiguous and provide prosodic cues to facilitate processing (cf., Clifton, Carlson, & Frazier, 2002; Grice, 1975). In line with previous ERP studies (e.g., Osterhout et al., 1994), the somewhat larger P600 effect for rejected than accepted sentences indicates that more difficult garden path sentences elicit increased P600 amplitudes more reliably and are more likely to be rejected. This relationship also holds in comparison to D.

Prosody–Syntax Mismatch Effects in Condition D

Condition D elicited a biphasic N400–P600 complex of ERP components. Our analyses ruled out the possibility that the N400 was a combination of other, prosody-related negativities, as it significantly differed in both size and topography from those. The N400 is likely to be similar to the one reported by Steinhauer et al. (1999) for German garden path sentences. In both studies, the N400 was elicited as soon as it became apparent that the current NP would not receive a proper thematic role. The difference between the studies is that in Steinhauer et al. (1999), the lack of this thematic role was indicated by an exclusively intransitive verb, whereas in the present study, two NPs

Table 5. P600 Effects in Conditions C (1300–1700 msec) and D (700–1300 msec), Relative to Splicing Point

<i>Source</i>	<i>df</i>	<i>P600 in D (700–1300 msec)</i>		<i>P600 in C (1300–1700 msec)</i>	
		<i>F</i>	<i>p</i>	<i>F</i>	<i>p</i>
<i>Lateral electrodes</i>		<i>Global ANOVA</i>		<i>Global ANOVA</i>	
Violation	1, 19	3.94	.07		
Prosody	1, 19	5.71	.03	9.90	.006
Violation × Lat	1, 19	10.49	.005	15.54	.001
Violation × Antpost	2, 38	8.94	.006		
Prosody × Hemi	1, 19	14.92	.001		
Prosody × Lat	1, 19			37.44	.0001
Prosody × Antpost	2, 38			8.23	.007
Viol × Pros × Lat	1, 19	11.45	.004		
Viol × Pros × Antpost	2, 38	7.78	.008		
Viol × Pros × Hemi	1, 19	21.02	.0003	14.96	.001
Viol × Pros × Lat × Antpost	2, 38			4.01	.03
Viol × Lat × Antpost × Hemi	2, 38	7.58	.005	3.66	.05
Viol × Pros × Lat × Antpost × Hemi	2, 38	7.59	.005	5.65	.02
<i>Midline electrodes</i>		<i>Global ANOVA (700–1300 msec)</i>		<i>Global ANOVA (1300–1700 msec)</i>	
Violation	1, 19	8.90	.008	5.89	.03
Prosody	1, 19	4.94	.04	22.49	.0001
Violation × Prosody	1, 19	4.44	.05		
Violation × Antpost	2, 38	7.88	.01		
Prosody × Antpost	2, 38	8.36	.002	7.05	.01
Viol × Pros × Antpost	2, 38			4.10	.05
<i>Lateral electrodes</i>		<i>Conditions A:D (700–1300 msec)</i>		<i>Conditions B:C (1300–1700 msec)</i>	
Violation	1, 19	7.50	.02		
Violation × Lat	1, 19	24.13	.0001	5.59	.03
Violation × Antpost	2, 38	10.51	.003	5.45	.03
Violation × Hemi	1, 19	12.18	.003	8.52	.01
Viol × Lat × Antpost × Hemi	2, 38	11.70	.0005		
<i>Midline electrodes</i>		<i>Conditions A:D (700–1300 msec)</i>		<i>Conditions B:C (1300–1700 msec)</i>	
Violation	1, 19	17.59	.0005		
Violation × Antpost	2, 38	4.24	.03	10.29	.003
Fz	1, 19	4.69	.04		
Cz	1, 19	21.43	.0002		
Pz	1, 19	19.32	.0003	7.83	.02

competed for the role to be assigned to the subject NP of the next clause before that verb was even present. In fact, both components were elicited even prior to the presentation of the third NP, and thus, must have been triggered by the second boundary. At this point, the second NP was prosodically separated from both the preceding verb and the subsequent clause.¹⁴ The N400 effect may be viewed as support for Bornkessel and Schlesewsky's (2006) proposal of an early processing stage during which so-called generalized semantic roles are being assigned to NPs in the absence of a verb, based on their prominence. The fact that no N400 was found in Condition C suggests that the lack of, rather than the necessity to revise, theta roles yields N400 components. This seems to be in line with previous findings of verb argument violations (e.g., Friederici & Frisch, 2000).

The N400 in Condition D was followed by a large positivity that we tentatively referred to as a P600. Given that the prosody–syntax mismatch condition involved two existing prosodic boundaries which prevented the second NP from being attached to the first verb (as object NP, due to Boundary 1) or the second verb (as subject NP, due to Boundary 2 and the presence of the third NP), a large P600 effect is not surprising. The amplitude of the P600 has repeatedly been demonstrated to increase with the severity of processing difficulties (e.g., Osterhout et al., 1994). However, the shape of the positive-going waveform in Condition D (particularly its steep onset and clear parietal peak) is also reminiscent of a P300 component known to reflect the updating of working memory (Donchin, 1981). The P300 is a domain-general (i.e., not specifically syntax-related) ERP component whose amplitude is positively correlated with the task relevance and inversely correlated with the probability of the stimulus or event that elicits it. There is a considerable literature reflecting a debate during the late 1990s on whether the P600, in general, should be viewed as a (delayed) P300 component (Coulson, King, & Kutas, 1998). Friederici, Mecklinger, Spencer, Steinhauer, and Donchin (2001) analyzed late positivities from a German garden path paradigm using PCA¹⁵ and argued that the P600 is not a monolithic ERP component and that the P300 may often be a subcomponent contributing to the P600. We believe that this is likely to be the case in Condition D as well. Given the severity of the syntax–prosody mismatch, the difficulty in resolving the problem (as it involves the deletion of a prosodic boundary), and the absence of any comprehension task that would have required the participants to actually perform a syntactic reanalysis, we believe the participants did not usually perform such a structural revision of the stimulus, but rather based their acceptability judgment on the mere detection of the anomaly. This would also explain the rather short response times for rejecting sentences in D (755 msec), which were not only some 200 msec faster compared to accepted trials in D (944 msec), but also more than 240 msec faster than rejections in the other three conditions (999–1080 msec, see Table 3). Together, the present pattern of garden path effects suggests that prosody does not only play a role in preventing or causing garden path ef-

fects, it also determines their *strength*. This supports Bader's (1998) claim that revision processes involving prosodic in addition to syntactic structures should be more difficult. In terms of syntactic processing alone, revisions from LC to EC (in C) and from EC to LC (in D) both involve changes of primary dominance relations (Gorrell, 1995), and can thus be viewed as comparable. However, they did differ in terms of prosodic reanalysis, supporting the BDH. This interpretation is also in line with the findings of Bögels et al. (2010), which demonstrated processing difficulty in sentences requiring the deletion of a superfluous prosodic boundary but not in sentences with a missing prosodic boundary. Greater processing difficulties related to deleting intonational phrase boundaries (e.g., in our Condition D) may be explained in various ways. First, any attempt to mentally undo the “positive evidence” of a boundary in the speech signal implies the listener's willingness to assume that the speaker mistakenly produced the salient boundary cues (compared to the more likely case of having missed an insufficient boundary marking in Condition C) (e.g., Clifton et al., 2002). Secondly, the extra time locally available at boundaries may help consolidate the initial syntactic analysis, potentially aided by conceptual–semantic interpretations of the established phrase (Selkirk, 1984). Recent evidence suggesting that the relative strength rather than the mere presence of boundaries may determine parsing decisions is compatible with the former (less localist) view. Whatever the exact mechanism at the boundary, only the prosodic phrasing seems to be reflected by the CPS (given that this component is elicited in absence of any lexical information). The strong similarities between data of the punctuation study (Steinhauer & Friederici, 2001) and the present results suggest that commas in written sentences and prosodic boundaries in speech have very similar effects, which are accounted for by the BDH. The boundary information seems to completely override any default structural parsing preferences such as the LC principle (see also Watson & Gibson, 2005). But why, then, did Kjelgaard and Speer (1999) find an overall advantage for LC, which appeared easier to process with conflicting prosody even compared to EC sentences with “neutral” prosody? We believe that differences in the stimulus materials may have played a role. Kjelgaard and Speer's LC garden path sentences with conflicting prosody all contained pronouns as the third NP [see example (a) below] and, therefore, allowed for an alternative revision [i.e., “left dislocation” as illustrated in (c) below] that would result in a grammatical structure *without* requiring a boundary deletion:

- (a) *When the maid cleans # the rooms they're immaculate* (garden path; prosody–syntax mismatch)
- (b) *When the maid cleans the rooms # they're immaculate* (supposed target structure after revision)
- (c) *When the maid cleans # the rooms—they're immaculate* (alternative structure after revision)

Interestingly, this alternative revision in (c) would not even differ in interpretation from the target LC structure.

In their EC garden path sentence, such an alternative revision is not possible, nor is it in other studies that employed lexical NPs and did not find the LC advantage (Steinhauer & Friederici, 2001; Walker et al., 2001).

The Closure Positive Shift

As in other ERP studies, the actual process of prosodic phrasing was reflected by the CPS. The CPS has previously been demonstrated for prosodic speech boundaries in German (Steinhauer et al., 1999), in Dutch (Kerkhofs et al., 2007) and, most recently, in Japanese (Wolff et al., 2008), Chinese (Li & Yang, 2009), and Korean (Hwang & Steinhauer, in press). Here we replicated this component for the first time in English sentences, thus adding cross-linguistic validity (see also Steinhauer, Abada, Pauker, Itzhak, & Baum, 2010, for CPS data in older people). Except for the second boundary in Condition D, where the CPS was superimposed by a large P600 component and could, therefore, not be analyzed independently, all three boundaries in Conditions A, B, and D displayed a strikingly consistent CPS pattern (Figure 6). In addition, ERP analyses aligned to the offset of the pre-boundary word yielded two important findings regarding the temporal fine structure of the CPS as well as preceding negativities. The CPS started right at (or even prior to) pause onset and shifted back toward the pre-CPS baseline after some 400 msec, potentially triggered by the onset of the next word. This is an important difference from the “music-CPS” found for phrase boundaries in melodies, whose onset was triggered by postboundary notes (Knoesche et al., 2005). This latency difference suggests that the perception of musical boundaries may require more contextual information. Further, at both the early and the late boundary, the CPS was preceded by negativities, the profile of which differed depending on the boundary position in the sentence. At the *first* boundary (in B and D), a central negativity was elicited during the auditory presentation of the preboundary VP1 (e.g., *is approaching*), reminiscent of the pre-CPS negativity reported by Bögels et al. (2010). The absence of this negativity in Conditions A and C strongly suggests that it was triggered by early prosodic cues carried by this VP which already marked the imminent boundary, that is, pre-boundary syllable lengthening and/or boundary tones. Future research will aim to specify the respective contribution of these two early acoustic boundary markers. In contrast, prior to the second boundary position in Condition A, a frontal negative shift built up, whose onset was already present near the onset of the pre-boundary NP2 (e.g., *the people*). This negative shift was (a) significantly more frontal than the one in B and D, and (b) it was also present in Condition C, which did not carry any acoustic boundary markers. Therefore, the negative shift in A and C must be viewed as reflecting a qualitatively different cognitive process than the pre-boundary negativity in B and D. We interpret this shift as an expectancy-related negativity for the following reasons. As the syntactic structures employed in this study required

a boundary either prior to (EC) or subsequent to (LC) the second NP, the lack of the first boundary in A and C may have led subjects to anticipate the presence of the second boundary. The fact that the amplitude was slightly larger in Condition A than in Condition C may argue for an additional contribution of prosodic cues (syllable lengthening and boundary tones) as well, comparable to Conditions B and D at the early boundary. After a “reset” of the frontal negative shift in A (Figure 6), the shape of the CPS was virtually the same as in B and D. This suggests that this reset and the CPS are additive and largely independent effects. Future research on the CPS will have to take these potential confounds into account; analyses across the entire sentence alone (rather than time-locked to the onset and offset of preboundary words or constituents) may not be able to tease these effects apart. Given that several previous studies investigating the CPS in speech (e.g., Bögels et al., 2010; Pannekamp et al., 2005; Steinhauer et al., 1999) also seem to show negativities preceding it, prosodic phrasing based on multiple boundary markers (including syllable lengthening, boundary tones and pauses) may normally include both negativities and CPS components. Unlike the expectancy-related negativity in A, the pre-CPS negativity in Conditions B and D may provide us with another valuable tool to study prosodic on-line processing, helping to tease apart the contribution of the various acoustic cues speakers use to mark boundaries.

Conclusion

Prosody-induced garden path effects in the ERP demonstrated that both the absence of a prosodic boundary and the presence of an inappropriate boundary result in processing difficulties that elicit typical ERP garden path effects (N400s and P600s). Our findings suggest that the LC principle guides listeners in the same way as readers, but that conflicting prosodic information immediately overrides this preference. Importantly, in line with previous reports on comma processing (e.g., Steinhauer & Friederici, 2001), structural revisions requiring the listener to insert a new boundary/comma (EC Condition C) appeared less difficult than those requiring the subsequent deletion of boundaries that were present in the original speech signal (LC Condition D). This finding lends strong support to the BDH and casts doubt on the generalizability of behavioral data, suggesting a general advantage of LC over EC sentences in cases of syntax–prosody mismatches. That is, at least in EC/LC ambiguities, the type of prosodic mismatch (missing vs. superfluous boundary) may be more important than the syntactic structure of the target sentence. Syntactic parsing and revision processes may further be modulated by an interaction between prosodic information and other factors, such as transitivity bias (see Itzhak et al., 2010).

The present study also replicated the CPS at prosodic boundaries in English sentences. Its early onset distinguishes the CPS at speech boundaries in adults not only from apparently similar positivities in infants (Männel &

Friederici, 2009) but also from previous “music-CPS” findings in skilled musicians (Knoesche et al., 2005), both of which occurred only after onset of the first word or tone following the boundary. Moreover, the CPS was preceded by negativities that seem to reflect the processing of early prosodic boundary markers realized on the preboundary word or phrase (syllabic lengthening, boundary tones).

Acknowledgments

We thank Meg Grant and Erin Vensel for assistance in stimulus recording. This work was supported by grants awarded to K. S. by the Canada Research Chair program and the Canada Foundation for Innovation (CRC/CFI; Project # 201876) and the Canadian Institutes of Health Research (CIHR; # MOP-74575), as well as a CIHR grant awarded to S. R. B. (# MOP-11290).

Reprint requests should be sent to Karsten Steinhauer, Faculty of Medicine, Centre for Research on Language, Mind and Brain, School of Communication Sciences & Disorders, McGill University, 1266 Pine Avenue West (Beatty Hall), Montreal, Quebec, Canada H3G-1A8, or via e-mail: karsten.steinhauer@mcgill.ca.

Notes

1. The acoustic markers of prosodic boundaries are realized as prefinal syllable lengthening, a boundary tone and a pause (see Cutler, Dahan, & van Donselaar, 1997).
2. By the same token, a prosodic pattern that conflicts with the sentence’s structure could mislead the parser toward an interpretation that is incompatible with upcoming information. The revision process in such a case is likely to be more difficult than a purely syntactic reanalysis because the prosodic pattern would have to be reanalyzed as well in order for the sentence to make sense (Bader, 1998).
3. Note that garden path effects in these German sentences (and in corresponding Dutch sentences studied by Bögels et al., 2010) can be accounted for by both the LC and the minimal attachment (MA) principles of the garden path model (Frazier, 1987). For reasons of simplicity, here we refer to the EC/LC ambiguity only (but see Bögels et al., 2010, and Steinhauer, 2003, for a discussion of the MA analysis).
4. To ensure that neither the behavioral nor the ERP effects could be attributed to audible splicing artifacts, we conducted a rating study including 13 subjects who judged both the original sentence conditions (A–D) and two newly created cross-spliced conditions (C₂ and D₂) that intentionally included splicing artifacts. Results: (i) Even unspliced conditions A + B were rated as “manipulated” in >20%; (ii) Conditions A + B vs. C + D differed by only ~13.5%, and all four conditions showed a comparable total range of acceptability (0–95%) across subjects; (iii) differences were at least partly due to participants’ off-line considerations/second thoughts (e.g., “if the sentence sounded weird, I concluded it was probably manipulated”), suggesting that subjects had major difficulties distinguishing between true splicing artifacts and the anomalous prosodic structure resulting from the cross-splicing (especially in D). In sum, we believe that there is little evidence for systematic cross-splicing artifacts in the materials used in our ERP study. This conclusion was also confirmed by a linguist with expertise in phonology/phonetics who listened to the full set of cross-spliced sentences used in our ERP study and found not a single audible splicing artifact.
5. Corresponding analyses for Conditions C and D were not necessary as they inherited this information from A and B during cross-splicing.

6. Note that the assignment of the levels of prosody to conditions changed after the splicing point: prior to the splicing point, A/C (no boundary) and B/D (with boundary) were contrasted, while after the splicing point, B/C (no boundary) and A/D (with boundary) were contrasted. This is illustrated in Table 2.
7. The laterality factor divided the electrodes into medial (F3/4, C3/4, P3/4) and lateral (F7/8, T3/4, T5/6) columns.
8. Five of the subjects did not accept any trial in either C or D, and thus, did not enter this analysis.
9. As this N400 in Condition D occurred after the splicing point, we will address it in more detail in the context of garden path effects.
10. With the –50 to +50 msec baseline, the CPS in Conditions B and D reached significance at 12 out of the 19 electrodes within the first 200 msec after offset of VP1.
11. In our ANOVA designs, this affected the levels of the factor prosody as follows: prior to the splicing point, A and C were assigned Level 1, and B and D were assigned Level 2. After the splicing point, Conditions A and D were assigned Level 1, and B and C were assigned Level 2. For the N400 effect in D, it was necessary to demonstrate that the negativity differed from both control conditions (A and B, respectively).
12. Note that a valid analysis of this kind, based on a sufficient number of trials, could be conducted only for Condition C, which was accepted in approximately half of the trials (whereas Conditions A and B were accepted, and D was rejected, in the vast majority of trials).
13. In addition, we also tested the influence of (in)transitivity biases (of the verbs used in these stimuli) on sentence parsing (see Itzhak et al., 2010).
14. It should be noted that a prosodic boundary followed by an NP at this position is, in itself, not necessarily indicative of an ungrammatical structure (as in *When a bear was approaching # the people # the dogs # and the sheep all looked up*). The early detection of the anomaly in our D condition may thus partly depend on the absence of corresponding correct sentences containing such boundaries. A follow-up ERP study that includes such sentences is currently underway in our lab to address these questions.
15. PCA = principal components analysis.

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 Academic Publishing.
- Beach, C. M. (1991). The interpretation of prosodic patterns at points of syntactic structure ambiguity: Evidence for cue trading relations. *Journal of Memory and Language*, 30, 644–663.
- Beckman, M. E. (1996). The parsing of prosody. *Language and Cognitive Processes*, 11, 17–67.
- Bögels, S., Schriefers, H., Vonk, W., Chwilla, D. J., & Kerkhofs, R. (2010). The interplay between prosody and syntax in sentence processing: The case of subject- and object-control verbs. *Journal of Cognitive Neuroscience*, 22, 1036–1053.
- Bornkessel, I., & Schlesewsky, M. (2006). The extended argument dependency model: A neurocognitive approach to sentence comprehension across languages. *Psychological Review*, 113, 787–821.
- Clifton, C., Carlson, K., & Frazier, L. (2002). Informative prosodic boundaries. *Language and Speech*, 45, 87–114.
- Coulson, S., King, J., & Kutas, M. (1998). Expect the unexpected: Event-related brain responses to morphosyntactic violations. *Language and Cognitive Processes*, 13, 21–58.
- Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, 40, 141–201.

- Donchin, E. (1981). Surprise!..Surprise? *Psychophysiology*, *18*, 493–513.
- Eckstein, K., & Friederici, A. D. (2006). It's early: Event-related potential evidence for initial interaction of syntax and prosody in speech comprehension. *Journal of Cognitive Neuroscience*, *18*, 1696–1711.
- 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., Carlson, K., & Clifton, C. (2006). Prosodic phrasing is central to language comprehension. *Trends in Cognitive Sciences*, *10*, 244–249.
- Frazier, L., Carminati, M. N., Cook, A. E., Majewski, H., & Rayner, K. (2006). Semantic evaluation of syntactic structure: Evidence from eye movements. *Cognition*, *99*, B53–B62.
- Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, *6*, 291–325.
- Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Neuropsychology*, *14*, 178–210.
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*, 78–84.
- 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*, 305–323.
- Gorrell, P. (1995). *Syntax and parsing*. Cambridge, UK: Cambridge University Press.
- Grice, P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics* (Vol. 3). New York: Academic Press.
- Hwang, H., & Steinhauer, K. (in press). Phrase length matters: The interplay between implicit prosody and syntax in Korean 'garden path' sentences. *Journal of Cognitive Neuroscience*.
- Itzhak, I., Pauker, E., Drury, J. E., Baum, S. R., & Steinhauer, K. (2010). Event-related potentials show online influence of lexical biases on prosodic processing. *NeuroReport*, *21*, 8–13.
- Kaan, E., & Swaab, T. Y. (2003). Repair, revision, and complexity in syntactic analysis: An electrophysiological differentiation. *Journal of Cognitive Neuroscience*, *15*, 98–110.
- Kerkhofs, R., Vonk, W., Schriefers, H., & Chwilla, D. (2007). Discourse, syntax, and prosody: The brain reveals an immediate interaction. *Journal of Cognitive Neuroscience*, *19*, 1421–1434.
- Kerkhofs, R., Vonk, W., Schriefers, H., & Chwilla, D. J. (2008). Sentence processing in the visual and auditory modality: Do comma and prosodic break have parallel functions? *Brain Research*, *1224*, 102–118.
- Kjelgaard, M. M., & Speer, S. R. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic ambiguity. *Journal of Memory and Language*, *40*, 153–194.
- Knoesche, T. R., Neuhaus, C., Hauelsen, J., Alter, K., Maess, B., Witte, O. W., et al. (2005). Perception of phrase structure in music. *Human Brain Mapping*, *24*, 259–273.
- Kutas, M., & Hillyard, S. A. (1980). Event-related brain potentials to semantically inappropriate and surprisingly large words. *Biological Psychology*, *11*, 99–116.
- Kutas, M., van Petten, C. K., & Kluender, R. (2006). Psycholinguistics electrified II. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 659–724). New York: Academic Press.
- Li, W., & Yang, Y. (2009). Perception of prosodic hierarchical boundaries in Mandarin Chinese sentences. *Neuroscience*, *158*, 1416–1425.
- Liu, B. L., Wang, Z. N., & Jin, Z. (2010). The effects of punctuation in Chinese sentence comprehension: An ERP study. *Journal of Neurolinguistics*, *23*, 66–80.
- Männel, C., & Friederici, A. D. (2009). Pauses and intonational phrasing: ERP studies in 5-month-old German infants and adults. *Journal of Cognitive Neuroscience*, *21*, 1988–2006.
- Marslen-Wilson, W. D., Tyler, L. K., Warren, P., Grenier, P., & Lee, C. S. (1992). Prosodic effects in minimal attachment. *Quarterly Journal of Experimental Psychology*, *45A*, 73–87.
- Mietz, A., Toepel, U., Ischebeck, A., & Alter, K. (2008). Inadequate and infrequent are not alike: ERPs to deviant prosodic patterns in spoken sentence comprehension. *Brain and Language*, *104*, 159–169.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: Edinburgh inventory. *Neuropsychologia*, *9*, 97–113.
- Osterhout, L., Holcomb, P., & Swinney, D. A. (1994). Brain potentials elicited by garden-path sentences: Evidence of the application of verb information during parsing. *Journal of Experimental Psychology*, *20*, 786–803.
- Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*, 785–806.
- Pannekamp, A., Toepel, U., Alter, K., Hahne, A., & Friederici, A. D. (2005). Prosody-driven processing: An event-related potential study. *Journal of Cognitive Neuroscience*, *17*, 407–421.
- Price, P. J., Ostendorf, M., Shattuck-Hufnagel, S., & Fong, C. (1991). The use of prosody in syntactic disambiguation. *Journal of Acoustical Society of America*, *90*, 2956–2970.
- Schafer, A. J., Speer, S. R., Warren, P., & White, S. D. (2000). Intonational disambiguation in sentence production and comprehension. *Journal of Psycholinguistics*, *29*, 169–182.
- Selkirk, E. (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.
- Snedeker, J., & Trueswell, J. (2003). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language*, *48*, 103–130.
- Staub, A. (2007). The parser doesn't ignore intransitivity, after all. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *33*, 550–569.
- Steinhauer, K. (2003). Electrophysiological correlates of prosody and punctuation. *Brain and Language*, *86*, 142–164.
- Steinhauer, K., Abada, S. H., Pauker, E., Itzhak, I., & Baum, S. R. (2010). Prosody-syntax interactions in aging: Event-related potentials reveal dissociations between on-line and off-line measures. *Neuroscience Letters*, *472*, 133–138.
- Steinhauer, K., Alter, K., & Friederici, A. D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech. *Nature Neuroscience*, *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 of prosodic phrasing in listeners and readers. *Journal of Psycholinguistic Research*, *30*, 267–295.
- Walker, J., Fongemie, K., & Daigle, T. (2001). Prosodic facilitation in the resolution of syntactic ambiguities in subjects with left and right hemisphere damage. *Brain and Language*, *78*, 169–196.
- Watson, D., & Gibson, E. (2005). Intonational phrasing and constituency in language production and comprehension. *Studia Linguistica*, *53*, 279–300.
- Wilson, M. P., & Garnsey, S. M. (2009). Making simple sentences hard: Verb bias effects in simple direct object sentences. *Journal of Memory and Language*, *60*, 368–392.
- Wolff, S., Schlesewsky, M., Hirotoni, M., & Bornkessel-Schlesewsky, I. (2008). The neural mechanisms of word order processing revisited: Electrophysiological evidence from Japanese. *Brain and Language*, *107*, 133–157.