

Pauses and Intonational Phrasing: ERP Studies in 5-month-old German Infants and Adults

Claudia Männel and Angela D. Friederici

Abstract

■ In language learning, infants are faced with the challenge of decomposing continuous speech into relevant units, such as syntactic clauses and words. Within the framework of prosodic bootstrapping, behavioral studies suggest infants approach this segmentation problem by relying on prosodic information, especially on acoustically marked intonational phrase boundaries (IPBs). In the current ERP study, we investigate processing of IPBs in 5-month-old infants by varying the acoustic cues signaling the IPB. In an experiment in which pitch variation, vowel lengthening, and pause cues are present (Experiment 1), 5-month-old German infants show an ERP obligatory response. This obligatory response signals lower level perceptual processing of acoustic cues that, however, disappear when no pause

cue is present (Experiment 2). This suggests that infants are sensitive to sentence internal pause, a cue that is relevant for the processing of IPBs. Given that German adults show both the obligatory components and the closure positive shift, a particular ERP component known to reflect the perception of IPBs, independent of the presence of a pause cue, the results of the current ERP study indicate clear developmental differences in intonational phrase processing. The comparison of our neurophysiological data from German-learning infants with behavioral data from English-learning infants furthermore suggests cross-linguistic differences in intonational phrase processing during infancy. These findings are discussed in the light of differences between the German and the English intonation systems. ■

INTRODUCTION

Since its inception in the seminal article by Gleitman and Wanner (1982), the concept of prosodic bootstrapping has received a lot of attention in the field of language acquisition (see also Gleitman, Gleitman, Landau, & Wanner, 1988; Morgan, Meier, & Newport, 1987; Peters, 1983). Gleitman and Wanner first proposed that acoustic cues in the speech input provide infants with cues to syntactic boundaries; the infant can bootstrap from the acoustic cues in the speech signal into the syntactic units defining the sentence structure and, thus, derive the basic syntactic rules governing the language input. Many studies have focused on identifying the acoustic cues provided in a given language and on determining the developmental stages during which infants encode and use those cues for speech segmentation.

Acoustic analyses of continuous speech show that syntactic boundaries often coincide with prosodic boundaries. Specifically, vowel lengthening, pitch change, and pause often combine to mark the edges of syntactic units, thus promoting, for example, the detection of clauses and phrases. Although not all syntactic boundaries are marked by prosodic boundaries, nearly every prosodic boundary is also a syntactic boundary.¹ It follows that prosodic bound-

aries may be, although not exhaustively (Gerken, Jusczyk, & Mandel, 1994), infants' first indicator for where syntactic boundaries occur in the speech stream (Gerken, 1996; Saffran, Newport, & Aslin, 1996).

According to the model of hierarchical prosodic structure (Selkirk, 1984; see also Nespor & Vogel, 1986), intonational phrases (IPs) are the largest prosodic constituents of an utterance and roughly correspond to syntactic clauses. IPs typically contain at least one nuclear accent and are characterized by the pitch contour course, that is, the lowering or rising of the contour toward the end of the phrase followed by a resetting of the pitch level at the beginning of the following phrase. Numerous studies demonstrate the existence of acoustic boundary cues in relation to clause and phrase boundaries. Specifically, at the right edge of syntactic phrases in syllables that precede the syntactic boundary, an increase in vowel length can be observed (e.g., Wightman, Shattuck-Hufnagel, Ostendorf, & Price, 1992; Beckman & Edwards, 1990; Cooper & Paccia-Cooper, 1980). In addition, the pitch contour signals the presence of clause boundaries by changes in pitch level (e.g., Ladd, 1988; Pierrehumbert, 1980). Finally, pauses at word boundaries that coincide with clause boundaries tend to be longer than pauses at word boundaries occurring elsewhere in an utterance (e.g., Scott, 1982; Cooper & Paccia-Cooper, 1980). These prosodic boundary cues are typically enhanced in infant-directed speech (Fisher & Tokura, 1996; Fernald et al., 1989).

Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

There is considerable behavioral evidence suggesting that very young infants are sensitive to the prosodic features of their native language. Several studies have shown that newborns readily distinguish their mother tongue from other languages based on its prosodic properties (e.g., Mehler et al., 1988). Also, from birth, infants seem to perceive some boundary information and can discriminate between identical syllable sequences that contain an intervening phrase boundary from those that do not (Christophe, Mehler, & Sebastián-Gallés, 2001; Christophe, Dupoux, Bertoncini, & Mehler, 1994). At around 4.5 months of age, infants begin to show sensitivity to specific prosodic markers in the continuous speech stream and prefer to listen to passages with artificial pauses inserted at clause boundaries over pauses inserted at other sentence positions (Jusczyk, Hohne, & Mandel, 1995; see also Hirsh-Pasek et al., 1987). At the phrasal level, the preference for pauses at phrasal over nonphrasal boundaries is evident at the age of 9 months; this has even been found for low-pass filtered speech, indicating that infants place a special emphasis on suprasegmental over segmental cues (Jusczyk et al., 1992). Further studies in English-learning infants have demonstrated their developing ability to recognize larger linguistic units in continuous speech based on prosodic boundary cues at both clause level (Seidl, 2007; Soderstrom, Nelson, & Jusczyk, 2005; Nazzi, Kemler Nelson, Jusczyk, & Jusczyk, 2000) and phrase level (Soderstrom, Seidl, Nelson, & Jusczyk, 2003). Infants at 6 months distinguish prosodically well-formed syntactic units from nonunits at clause level but not at phrase level (Seidl, 2007; Soderstrom et al., 2003, 2005; Nazzi et al., 2000), whereas infants at 9 months show this distinction even at phrase level (Soderstrom et al., 2003). Thus, the existing data suggest that infants learning English perceive larger linguistic units, like clauses, earlier in continuous speech than smaller units, such as syntactic phrases.

The recognition of prosodic phrase boundaries may also facilitate the detection of words because phrase boundaries and word boundaries coincide at least at the right edge of the phrase boundary. Infants seem not to segment words from fluent speech until they can recognize larger units in the speech stream because they do not start to extract monosyllabic words from fluent speech before 6–7.5 months (Jusczyk & Aslin, 1995) and bisyllabic words before 7.5 months (Nazzi, Dilley, Jusczyk, Shattuck-Hufnagel, & Jusczyk, 2005; Johnson & Jusczyk, 2001; Jusczyk, Houston, & Newsome, 1999; Morgan & Saffran, 1995). Gout, Christophe, and Morgan (2004) demonstrated that for 10-month-old infants, word detection was facilitated when words occurred at phrase boundary positions (see also Seidl & Johnson, 2007). Thus, the analysis and the segmentation of smaller, syntactically and lexically relevant units occur later during an infant's language development and are likely to be facilitated by bootstrapping from acoustically marked larger units in the speech signal.

It is still an open question which acoustic cues are required to trigger the perception of prosodic boundaries

in a given language at different developmental stages. Behavioral studies in English suggest that adults rely more heavily on both pitch changes and preboundary lengthening than on pausing (Aasland & Baum, 2003; Wightman et al., 1992; Scott, 1982; Streeter, 1978). For infants learning English, Seidl (2007) found that pitch plays an important role in phrase boundary perception. However, pitch alone was not sufficient for triggering phrase boundary perception; it had to co-occur with either preboundary lengthening or with pausing. Similarly, Gerken et al. (1994) found that English-learning infants required a combination of both pitch change and preboundary lengthening to identify phrase boundaries. Thus, there is some evidence, at least for English stimulus material, that infants and adults weight specific boundary cues differently in the perception of phrasal units. In adults, only the manipulation of boundary length affected speech segmentation, whereas in infants, either boundary lengthening or pausing had to co-occur with particular variations in pitch patterns.

Given the potential significance of the prosodic processing abilities for later stages of lexical-semantic and syntactic acquisition, it is of great interest to investigate the principles underlying the infants' early perception of prosodic cues that signal IPs. Although a number of behavioral studies have contributed to the description of the developmental course of infants' sensitivity to IPs, its underlying neurophysiological basis remains unknown. However, the neurophysiology can be explored using ERP, an investigational method that provides an on-line measure of the ongoing stimulus processing at different levels, from perceptual to cognitive processing (see, e.g., Friederici & Thierry, 2008). Early components in the ERP, occurring between 100 and 200 msec, are known to reflect sensory-perceptual levels of processing, whereas later ERP components reflect cognitive processes. In the late processing domain, a specific electrophysiological component has been found, that is, a positive shift in the ERP that occurs in correspondence to intonational phrase boundaries (IPBs) in adults (Steinhauer, Alter, & Friederici, 1999; see also Kerkhofs, Vonk, Schriefers, & Chwilla, 2007; Pannekamp, Toepel, Alter, Hahne, & Friederici, 2005). This positive shift has been interpreted as indicating the closure of major prosodic phrases by phrase boundaries and is therefore called *closure positive shift* (CPS). ERP measures allow for identification of processing levels involved in principle and may also do so specifically for the processing of IPBs in infants.

In addition to investigating the neurophysiological correlates of infants' prosodic processing, it is of special interest to target German-learning infants' processing abilities because most of the behavioral studies on phrase and clause segmentation have been carried out in English. English and German intonation systems differ with respect to their particular acoustic realization of IPs. In English, IPs are predominantly marked by their functional pitch contour, whereas in German, length and loudness seem to be similarly important (Markus, 2006; Delattre, 1966).

Given the behavioral evidence for English-learning infants' sensitivity to prosodic phrasing, the present study has two aims. First, the ERP experiments will probe the ability of 5-month-old German-learning infants to process IPs. Here, we will specify the role of the pause as one of the three acoustic cues signaling prosodic boundaries. Using ERPs, we will determine whether infant responses to IPBs are primarily attributable to lower level, automatic processing of acoustic stimulus features, being reflected in obligatory components² such as the N1–P2 complex in adults (Crowley & Colrain, 2004; Näätänen & Picton, 1987), or whether they indicate higher level cognitive processing, being reflected in the CPS (Steinhauer et al., 1999). Second, the comparison of the infant and the adult ERP data collected in response to identical stimulus material will provide insight into developmental changes in IP processing. Specifically, we will consider the nature of the underlying processes and the potential developmental differences in the weighting of specific boundary cues.

METHODS

Subjects

Infants were recruited from the Infant Database of the Max Planck Institute for Human Cognitive and Brain Sciences in Leipzig and the Database of the German Language Development Study in Berlin. All participants were healthy, full-term 5-month-olds without any known hearing deficits or neurological problems. All infants came from monolingual German families. Before their participation, parental informed consent was obtained for all infants. The data of about one fourth of all infants tested were excluded because they contained disproportionate numbers of artifacts caused by exaggerated movement, perspiration (skin potentials), or crying. In Experiment 1, the final sample consisted of 34 infants (19 female; mean age = 22.1 weeks, range = 21–23 weeks); in Experiment 2, the final sample also comprised 34 infants (17 female; mean age = 22.2 weeks, range = 21–23 weeks).

All adult participants were students from the University of Leipzig who spoke German natively and were right-handed, as assessed by a German version of the Edinburgh Handedness Inventory (Oldfield, 1971). These subjects reported having no known hearing deficits or neurological problems. Experiment 1 involved 27 participants (14 female; mean age = 24.4 years, range = 18–30 years) and Experiment 2 included 28 participants (13 female; mean age = 24.5 years, range = 19–29 years).

Stimuli

Subjects were presented with either the naturally spoken sentences (Experiment 1) or the same sentences for which the pause parameter was neutralized between con-

ditions (Experiment 2). All sentences either contained or lacked an IPB, thus differing in their intonational realization; this was because the underlying syntactic structures differed dependent on the valence of the second verb. Sentences with IPB contained two IPs divided by an IPB, whereas sentences without IPB only consisted of one IP and, consequently, lacked an IPB (Table 1). Sentences with IPB contained a transitive verb at the end. For the example *Tommi verspricht Papa zu helfen* (Tommi promises to help papa), *zu helfen* (to help) is the transitive verb and it must be accompanied by a noun phrase, which in this case is *Papa* (direct object). As a consequence, an IPB occurred at the first verb, marking a first syntactic phrase (*Tommi verspricht*/Tommi promises) that was followed by a second phrase (*Papa zu helfen*/to help papa). In contrast, sentences without IPB ended with an intransitive verb; for example, *Tommi verspricht Papa zu schlafen* (Tommi promises papa to sleep) was followed by *zu schlafen* (to sleep). Here, the noun phrase *Papa* is the indirect object of the first verb *verspricht* (promises), and therefore sentences of this type did not contain a sentence-internal IPB.

The sentences were designed with each sentence pair being identical in wording up to the second verb, which could be either transitive or intransitive. The identical word order in German actually imposes a structural ambiguity on the listener's parsing process because the syntactic role of the noun phrase, that is, direct object or indirect object, is not resolved until the occurrence of the second verb at the end of the sentence, that is, the transitive verb or the intransitive verb.³ However, the syntactic structure can be easily determined before the appearance of the second verb because the prosodic parameters vary significantly between both sentence types from the sentence onset, with the prosodic contour signaling the respective syntactic units (see Acoustic parameters for naturally spoken sentences section).

All 100 sentences (50 sentences of each type) were produced in a soundproof chamber by a trained female speaker, who was instructed to speak in an infant-directed manner. After recording, sentences were digitized (44.1 kHz/16-bit sampling rate, mono) and normalized in amplitude to 70%.

Table 1. Stimulus Examples (with Literal Translations)

(1) Example sentence with IPB (two IPs)

[*Tommi verspricht*]_{IP 1} [*Papa zu helfen*]_{IP 2}

[Tommi promises]_{IP 1} [papa to help]_{IP 2}

(2) Example sentence without IPB (one IP)

[*Tommi verspricht Papa zu schlafen*]_{IP 1}

[Tommi promises papa to sleep]_{IP 1}

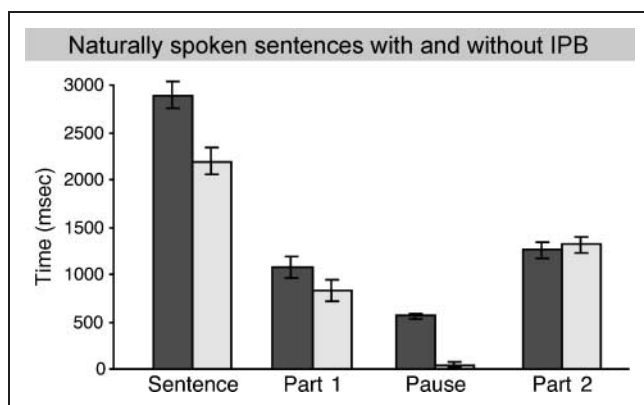


Figure 1. Naturally spoken sentences: duration parameters of sentences, sentence parts 1 and 2, and pauses for sentences with IPB (dark gray) and sentences without IPB (light gray; error bars indicate *SD*).

By using sentence material produced in an infant-directed manner, we ensured infants were presented with the characteristic language input they are usually exposed to in their natural environment. Infant-directed speech (also referred to as *motherese* or *baby talk*) is characterized by simplified structure, slow speech rate, and exaggerated prosodic features, for example, greater pitch variability and longer pauses (Fernald et al., 1989). In the current study, using infant-directed speech was important (a) for attracting the infants' attention to the stimulus materials and (b) for ensuring an effective experimental manipulation by presenting prosodic boundary features in a manner readily perceived by the infants (Cooper & Aslin, 1990; Fernald, 1985).

Acoustic Parameters for Naturally Spoken Sentences (Experiment 1)

In Experiment 1, naturally spoken short sentences were presented. Acoustic analyses (Adobe Audition 1.0, Adobe Systems, San Jose, CA; Praat 4.4.18, University of Amsterdam, Netherlands) revealed clear differences between sentences with IPB and sentences without IPB regarding their durational and tonal characteristics.

For the analysis, both sentence types were divided into three sections of interest by markers that were manually set in each audio file: (a) first sentence part *Tommi verspricht* (Tommi promises), (b) pause (following the first verb), and (c) second sentence part *Papa zu schlafen/helfen* (papa to sleep/to help papa). As can be seen in Figure 1, the first sentence part is significantly longer for sentences with IPB than for sentences without IPB (independent two-sampled *t* test), $t(98) = -10.78, p \leq .01$. This difference is mainly accounted for by the significantly longer syllable at the end of the first part in sentences with IPB, $t(98) = -22.99, p \leq .01$. In contrast to sentences without IPB, sentences with IPB included a rise in pitch at the end of the first sentence part, $t(98) = -36.22, p \leq .01$ (for the F0 contour from two example sentences, see Figure 2). Furthermore, the pause following the first sentence part is significantly longer, $t(98) = -106.38, p \leq .01$, for sentences with IPB (560 msec, $SD = 26$ msec) than for sentences without IPB (45 msec, $SD = 22$ msec). As a result, the overall sentence length differs significantly between conditions, $t(98) = -24.71, p \leq .01$; sentences with IPB averaged 2890 msec ($SD = 136$ msec) and sentences without IPB averaged 2200 msec ($SD = 145$ msec). In sum, the acoustic analyses confirmed that sentences with IPB indeed contained an IPB at the end

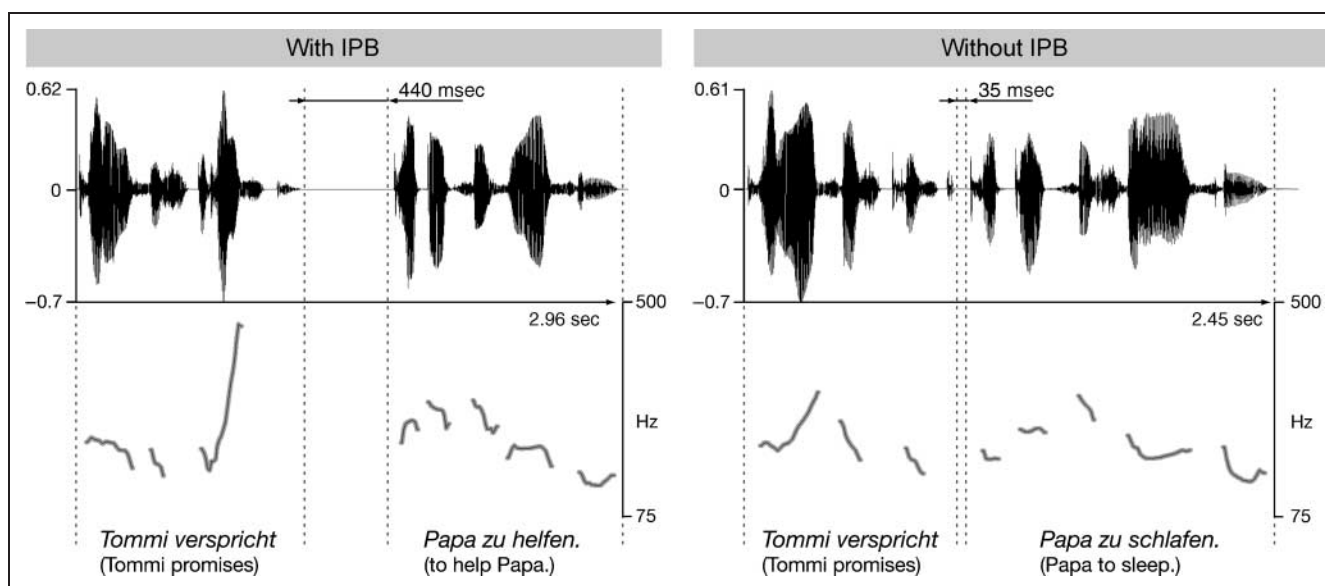


Figure 2. Naturally spoken sentences: waveform (normalized values) and pitch track (F0 contour in Hz) for examples of sentences with IPB (left) and sentences without IPB (right).

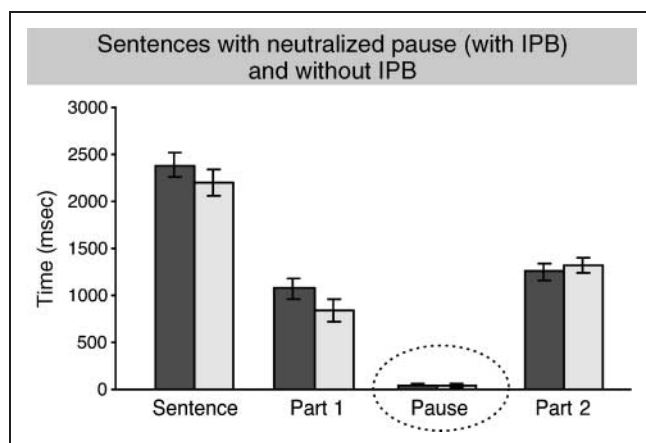


Figure 3. Sentences with neutralized pause: duration parameters of sentences, sentence parts 1 and 2, and pauses for sentences with IPB (dark gray) and sentences without IPB (light gray; error bars indicate *SD*). Note that except for pause length, all other duration parameters remained the same as in Experiment 1.

of the first sentence part, whereas sentences without IPB did not. The IPB was acoustically realized by an increase in pitch level, a lengthening of the last syllable, and a protracted pause following a phrase-final increase in pitch.

Acoustic Parameters for Sentences with Neutralized Pause (Experiment 2)

In Experiment 2, sentences were created from the naturally spoken sentences by neutralizing differences in pause length between conditions. This was done by removing the pause following the boundary tone in the sentence condition containing the IPB (see Figure 3), hereafter called *sentences with neutralized pause* (with

IPB). We deleted about 500 msec of silence, preserving voice offset and onset for the adjacent words, thus ensuring the naturalness of the new material by retaining the brief pause that typically occurs between words in sentences. The remaining pauses did not significantly differ from the pauses at the same sentence position in sentences without IPB, $t(98) = -1.03$, $p = .31$. Apart from shortening the pauses in the sentences with neutralized pause (with IPB), all other acoustic features remained the same for both sentence types (see Figure 4). As a result of pause deletion, the sentences with IPB were shorter than in Experiment 1, but there was still a significant difference in total duration between the IPB and without IPB conditions, $t(98) = -6.65$, $p \leq .01$. The mean length for sentences with neutralized pause (with IPB) was 2380 msec ($SD = 132$ msec), whereas for sentences without IPB it was 2200 msec ($SD = 145$ msec).

Procedure

During the experimental procedure, infants sat on their parent's lap or in a car seat in an electrically shielded and sound-attenuated testing booth. The sentence material was delivered via a loudspeaker, controlled by ERTS software (BeriSoft Corp., Frankfurt, Germany). Although the infants were listening to the sentences, a visual distraction (silent baby video or puppet) was presented on demand to keep them entertained. Adults passively listened to the sentences while keeping their eyes fixated on the screen. They were asked to refrain from blinking during sentence presentation.

Fifty sentences for each condition, with IPB and without IPB, were presented in a pseudorandom order in blocks of 25 trials. Subjects received one of eight lists,

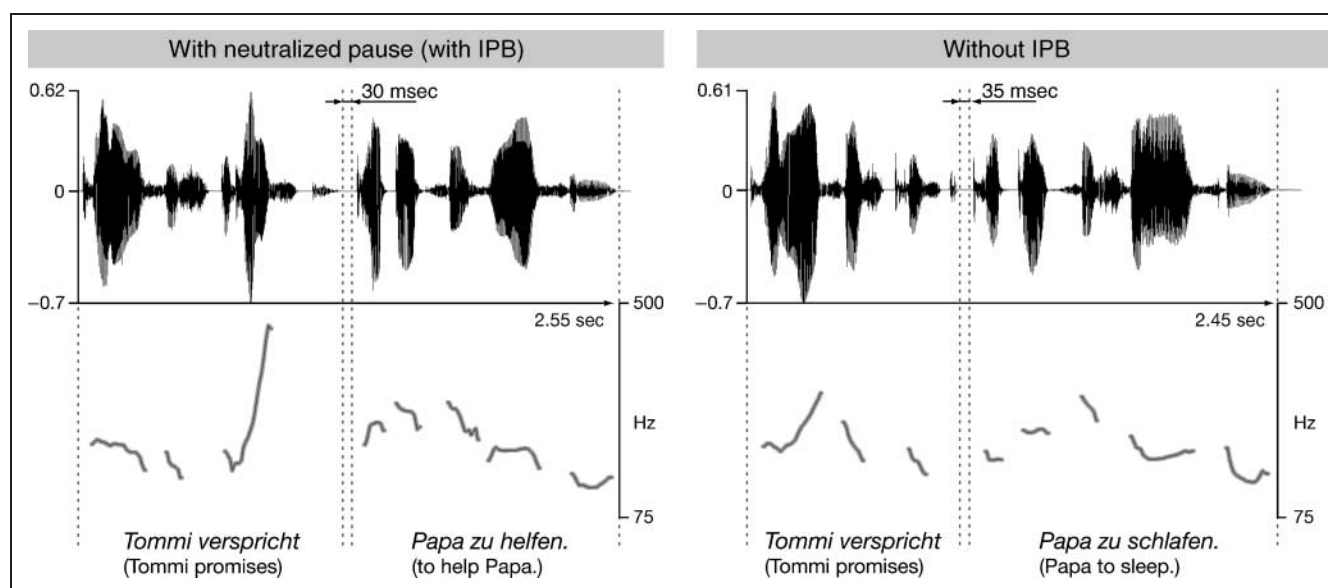


Figure 4. Sentences with neutralized pause: waveform (normalized values) and pitch track (F0 contour in Hz) for examples of sentences with IPB (left) and sentences without IPB (right).

in which the order of the sentences was randomized. Essentially, a sentence (2–3 sec) was presented, followed by an ISI of 1.5 sec. In total, Experiment 1 lasted about 8 min, whereas Experiment 2 lasted about 7 min due to the pause neutralization.

EEG Recordings

The EEG was continuously recorded from 23 Ag/AgCl ring electrodes attached to an elastic cap. The electrodes were positioned across the scalp according to the 10–10 system (Chatrian, Lettich, & Nelson, 1988): F7, F3, FZ, F4, F8, FC3, FC4, T7, C3, CZ, C4, T8, CP5, CP6, P7, P3, PZ, P4, P8, O1, O2, M1, and M2. In addition, an EOG was recorded from two electrodes at the outer canthi of both eyes (horizontal EOG) and from single electrodes on the infraorbital and supraorbital ridges of the right eye (vertical EOG). The recordings were referenced to CZ, whereas an additional electrode served as common ground (placed at FP1 for infants and on the sternum for adults). For infants, electrode impedances were in most cases kept below 10 k Ω (at least below 20 k Ω), and for adults in all cases, it was below 5 k Ω . The EEG signal was amplified with a gain of 20, using a PORT-32/MREFA (Twente Medical Systems International, Oldenzaal, Netherlands), with an input impedance of 10¹² Ω . The EEG data were digitized on-line at a rate of 250 Hz (AD converter with a 22-bit digital filter from DC to 125 Hz) and stored on a hard drive for further analyses.

Data Processing and Analysis

Off-line, the EEG data were processed using the EEP 3.3 software package (Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany). At first, the EEG data were algebraically re-referenced to the average of both mastoids (M1 and M2). To remove very slow drifts and muscle artifacts from the EEG, a digital band-pass filter ranging from 0.2 to 20 Hz (–3 dB cutoff frequencies of 0.25 and 19.91 Hz) was applied. Time segments of 3500 msec, relative to sentence onset, were extracted from the continuous EEG signal and adjusted to a prestimulus baseline of 200 msec.

In infants, EEG responses exceeding an *SD* of 70 μ V in a sliding window of 500 msec were considered invalid and were excluded. In addition, manual screening of all trials ensured the exclusion of artifacts not detected by the automatic rejection criterion (e.g., slow drifts). The remaining trials were separately averaged for sentences with and without IPB for each subject. At least 12 artifact-free trials were required per condition for an individual average to enter the final sample. The resulting mean number of averaged trials across subjects was 23 (*SD* = 9) in Experiment 1 and 25 (*SD* = 9) in Experiment 2; these numbers did not differ between conditions.

In adults, EEG epochs were individually checked for eye blinks and eye movements and corrected by a computer

algorithm (implemented in EEP 3.3). All other artifacts were detected manually. In Experiment 1, the resulting mean number of averaged trials across subjects was 46 (*SD* = 3) for sentences with IPB and 45 (*SD* = 4) for sentences without IPB, and in Experiment 2, it was 44 (*SD* = 4) for sentences with IPB and 43 (*SD* = 5) for sentences without IPB. In both experiments, trial numbers did not significantly differ between conditions.

Statistical analyses were performed separately for mean amplitudes on midline and lateral electrode sites. For midline sites (FZ, CZ, and PZ), a two-way ANOVA was computed with the factors Condition (with IPB and without IPB) and Region (anterior, central, and posterior). For lateral sites, six ROIs were created by combining hemisphere (left and right) and region (anterior, central, and posterior) information. Each of the six ROIs covered three electrode sites: left anterior (F7, F3, and FC3), right anterior (F8, F4, and FC4), left central (T7, C3, and CP5), right central (T8, C4, and CP6), left posterior (P7, P3, and O1), and right posterior (P8, P4, and O2). Subsequently, a three-way ANOVA was performed with the factors Condition (with IPB and without IPB), Region (anterior, central, and posterior), and Hemisphere (left and right). Significant interactions involving the factor Condition were further analyzed using one-way ANOVAs for regions, hemispheres, and sites/ROIs, respectively. All ANOVAs were calculated for mean amplitudes in defined time windows (TWs) or mean peak amplitude measures (see Results section); the Greenhouse–Geisser correction (Greenhouse & Geisser, 1959) was applied when there was more than one degree of freedom (*df*) in the numerator to account for potential violations of sphericity.

RESULTS

Naturally Spoken Sentences (Experiment 1)

For the initial statistical analysis of the ERP data, ANOVAs were performed with the factors Condition, Region, and Hemisphere across TWs of 500 msec, relative to sentence onset. Only significant amplitude differences involving the factor Condition are reported here.

For adults, Figure 5 (left panel) displays the ERP responses to sentences with IPB and sentences without IPB relative to sentence onset. The ERP response to sentences with IPB, but not to sentences without IPB, revealed a positive shift starting at around 1000 msec after sentence onset that lasted for about 1000 msec. This effect was validated by the results of ANOVAs performed across TWs of 500 msec, starting at sentence onset. As the results in Table 2 demonstrate, main effects of condition were observed for the TWs 1000–1500 and 1500–2000 msec.⁴ Although for both TWs the Condition \times Region interaction effects suggested topographical differences, the condition effect was broadly distributed over the scalp. Additional effects involving condition were observed before and after the time interval of 1000–2000 msec;

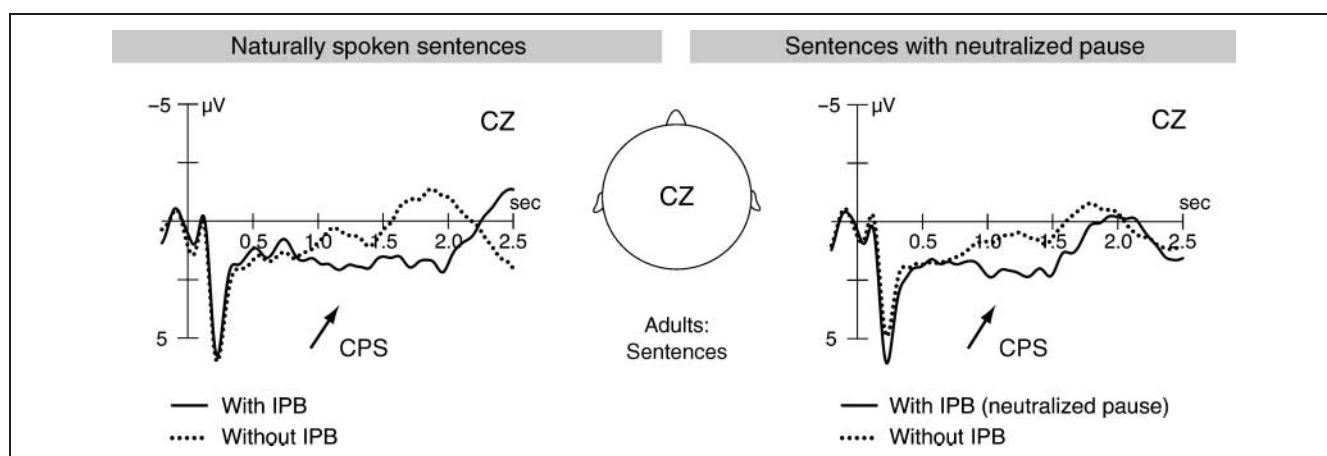


Figure 5. Adults: grand average ERPs (displayed at CZ) to naturally spoken sentences (left) and sentences with neutralized pause (right) for sentences with IPB (solid line) and without IPB (dotted line).

however, subsequent one-way ANOVAs for regions and hemispheres did not reveal any condition differences. In sum, the statistical analyses revealed processing differences between sentences with IPB and sentences without IPB that were apparent in a positive shift in the ERP in response to sentences containing IPBs.

In infants, the ERP responses to sentences with IPB showed a clear positive shift starting at around 2000 msec after sentence onset that lasted for about 1000 msec and which did not occur for sentences without IPB (Figure 6, left panel). Accordingly, ANOVAs revealed a main effect

of condition for the TW 2000–2500 msec at midline electrodes, $F(1, 33) = 14.63, p \leq .01$, and lateral ROIs, $F(1, 33) = 11.05, p \leq .01$. In addition, a Condition \times Region interaction was observed at lateral sites, $F(2, 66) = 9.2, p \leq .01$. The main effect condition continued for the subsequent TW 2500–3000 msec at both midline, $F(1, 33) = 10.79, p \leq .01$, and lateral sites, $F(1, 33) = 8.77, p \leq .01$. For the same TW, a Condition \times Region interaction was again found at lateral ROI positions, $F(2, 66) = 5.09, p \leq .05$. When testing the interaction in both TWs, one-way ANOVAs performed for each region revealed main effects

Table 2. Adults—Naturally Spoken Sentences: Significant Effects of ANOVAs for the 0- to 2500-msec Latency Range Relative to Sentence Onset

TW (msec)	Lateral ROIs			Midline Sites		
	Effect	df	F	Effect	df	F
500–1000	Cond \times Reg	2, 52	6.28*			
1000–1500	Cond	1, 26	14.9**	Cond	1, 26	15.49**
	Cond \times Reg	2, 52	8.65**	Cond \times Reg	2, 52	3.95*
	Ant	1, 26	4.54*	FZ	1, 26	6.16*
	Cen	1, 26	8.8**	CZ	1, 26	12.23**
	Post	1, 26	26.43**	PZ	1, 26	17.1**
1500–2000	Cond	1, 26	39.26**	Cond	1, 26	25.24**
	Cond \times Reg	2, 52	6.15*	Cond \times Reg	2, 52	10.77**
	Ant	1, 26	15.19**	FZ	1, 26	10.25**
	Cen	1, 26	53.66**	CZ	1, 26	27.56**
	Post	1, 26	46.67**	PZ	1, 26	30.8**
2000–2500	Cond \times Hem	1, 26	10.99**			

Cond = condition; Reg = region; Hem = hemisphere; Ant = anterior; Cen = central; Post = posterior.

* $p \leq .05$.

** $p \leq .01$.

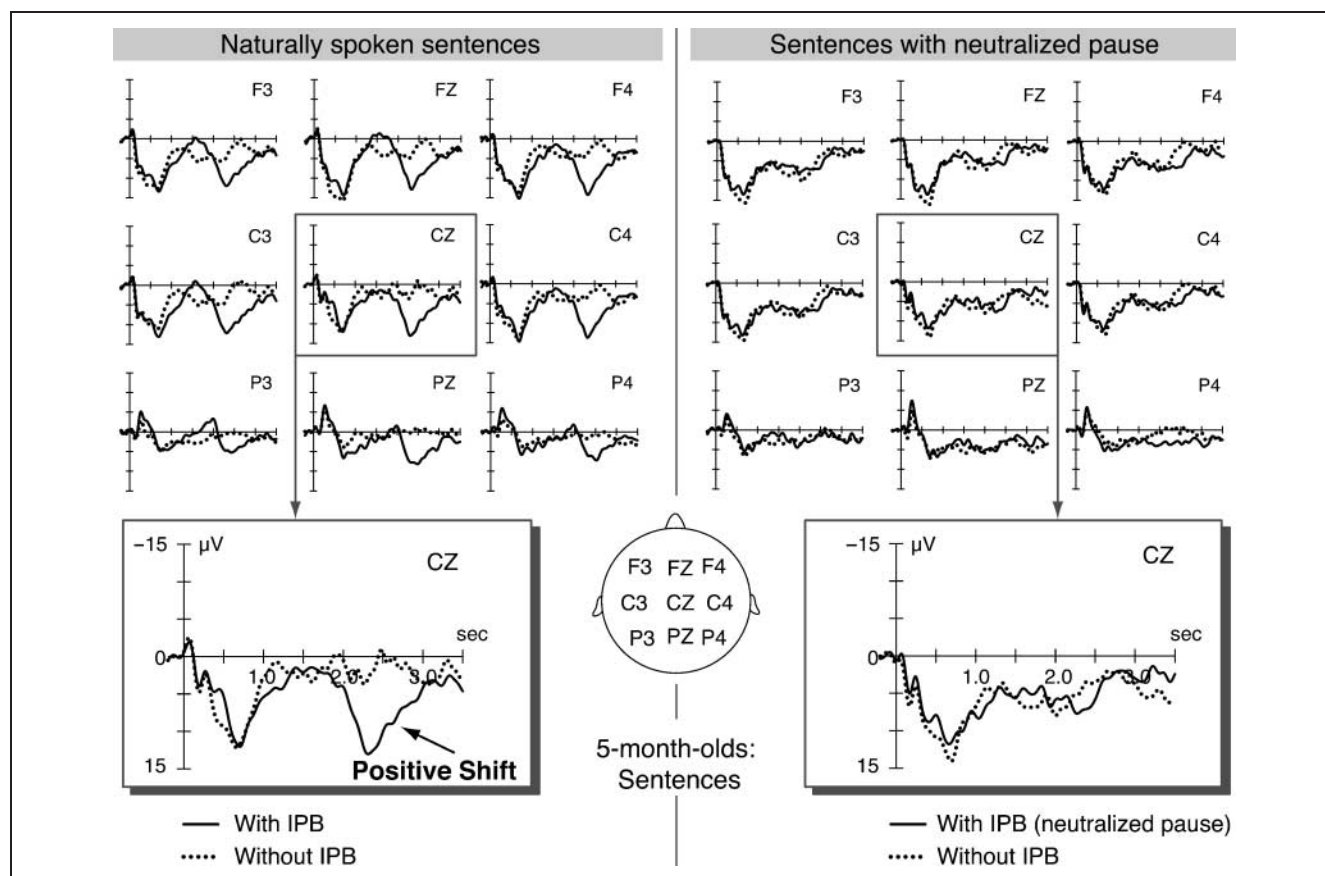


Figure 6. Five-month-olds: grand average ERPs to naturally spoken sentences (left) and sentences with neutralized pause (right) for sentences with IPB (solid line) and without IPB (dotted line). For naturally spoken sentences, ERPs to sentences with IPB display a positive shift, starting at about 2000 msec relative to sentence onset. For sentences with neutralized pause, ERPs to sentences with IPB (without pause) do not display a positive shift.

of condition at anterior, TW 2000–2500 msec, $F(1, 33) = 13.82$, $p \leq .01$, and TW 2500–3000 msec, $F(1, 33) = 10.94$, $p \leq .01$, and central regions, TW 2000–2500 msec, $F(1, 33) = 14.67$, $p \leq .01$, and TW 2500–3000 msec, $F(1, 33) = 11.37$, $p \leq .01$. Thus, the statistical analyses revealed the occurrence of a positive shift in the sentences with IPB, predominantly at anterior and central sites, whereas none was observed in the sentences without IPB.

Sentences with Neutralized Pause (Experiment 2)

As in Experiment 1, ANOVAs were performed with the factors Condition, Region, and Hemisphere across TWs of 500 msec relative to sentence onset.

For adults, the ERP responses to sentences with neutralized pause (with IPB) and sentences without IPB are given in Figure 5 (right panel). Similar to the ERP pattern displayed in Figure 5 (left panel), sentences with neutralized pause (with IPB) evoked a positive shift, whereas no such shift was observed for the sentences without IPB. Again, ANOVAs delivered effects involving the factor Condition for the TWs 1000–1500 msec and 1500–2000 msec at lateral ROIs and midline sites (Table 3). Thus, statistical analyses revealed processing differences between both sentence

types, indicated by a positive shift in the sentences with IPB, even in the absence of the pause boundary cue.

For infants, the ERP responses to sentences with neutralized pause (with IPB) and sentences without IPB are displayed in Figure 6 (right panel). In contrast to the ERP pattern displayed in Figure 6 (left panel), there is no positive shift in the sentences with IPB within the reported time range. Accordingly, the ANOVAs did not deliver any significant condition differences in the seven TWs previously analyzed. However, the ERP waveforms indicated condition differences, starting later and being of a shorter duration than the effects observed in Experiment 1. As the waveforms suggest, the ANOVA revealed a main effect of condition at lateral sites in the newly defined TW 2200–2700 msec, $F(1, 33) = 10.49$, $p \leq .01$.

The condition difference can be accounted for by either of two explanations. This difference may reflect a true condition effect, which was characterized by a longer latency and a shorter duration in comparison to that of Experiment 1. Alternatively, this effect may have resulted from the different sentence lengths; sentences with neutralized pause (with IPB) were about 200 msec longer than sentences without IPB. Thus, condition differences might stem from the respective termination of stimulus

Table 3. Adults—Sentences with Neutralized Pause: Significant Effects of ANOVAs for the 0- to 2500-msec Latency Range Relative to Sentence Onset

TW (msec)	Lateral ROIs			Midline Sites		
	Effect	df	F	Effect	df	F
1000–1500	Cond	1, 27	10.59**	Cond	1, 27	14.95**
	Cond × Reg	2, 54	7.47**	Cond × Reg	2, 54	5.44*
	Ant	1, 27	16.69**	FZ	1, 27	21.0**
	Cen	1, 27	9.58**	CZ	1, 27	16.67**
1500–2000	Cond	1, 27	11.84**	PZ	1, 27	5.54**
	Cond × Reg	2, 54	14.31**	Cond	1, 27	8.01**
2000–2500	Cond × Reg	2, 54	14.31**	Cond × Reg	2, 54	14.16**
	Post	1, 27	8.76**	PZ	1, 27	6.29*

Cond = condition; Reg = region; Ant = anterior; Cen = central; Post = posterior.

* $p \leq .05$.

** $p \leq .01$.

processing in each sentence condition that accordingly stops earlier in the sentences with neutralized pause (with IPB) than in the sentences without IPB. To determine which of these hypotheses better explained this phenomenon, we minimized the duration differences between the two sentence types by shifting the ERP analysis window to the onset of the second sentence part (i.e., pause offset). If the condition difference remains, then there must be a true condition effect, and if not, the sentence duration must be the critical feature. Because the ANOVA performed for three TWs across 500 msec revealed no condition effect, we conclude that the condition difference observed in Experiment 2 is primarily attributable to the different sentence lengths and does not derive from the presence of the IPB in the sentences with neutralized pause.

Positive Shift in the Infant ERP—CPS or Obligatory Onset Response

Given the resemblance of the positive shift observed in the naturally spoken sentences with IPB (Experiment 1) and the pronounced obligatory positive wave at their corresponding sentence onsets (especially at anterior and central sites), we further investigated this positive deflection following the IPB. Each sentence onset or phrase onset following a pause evokes an obligatory response that signals speech input after a period of silence. In infants, this obligatory response is reflected in positive waves at frontal and central electrode sites (P150 and P350; see Kushnerenko et al., 2002). It follows that the positive deflection following the IPB may reflect (a) a true phrasing effect comparable to the CPS observed in adults, (b) an obligatory response to the onset of the second part of the sentence, or (c) a combination of both these effects (for a methodological distinction between

the CPS and the adult obligatory component P2, see also Steinhauer, 2003). To further examine the positive shift observed in sentences with IPB, we directly compared the ERPs evoked by the sentence onset and the onset of the second noun phrase (henceforth called NP2 onset) by looking for response disparities between sentence Part 1 (*Tommi verspricht*/Tommi promises) and sentence Part 2 (*Papa zu helfen*/to help papa). If, in addition to the obligatory onset response, the observed positive deflection also includes a CPS, this should be apparent in at least one of the following ways: (a) a difference in amplitude (a larger amplitude would be expected for NP2 than for sentence onsets due to additive effects), (b) a difference in scalp distribution (produced by amplitude differences at posterior sites for the CPS), or (c) a difference in the appearance of the waveforms (e.g., overlap at posterior sites for the CPS). For this analysis, we averaged EEG epochs of 1500 msec relative to their corresponding onsets, adjusted to a prestimulus baseline of 50 msec.⁵ ANOVAs were performed across TWs of 500 msec with the factors Onset Type, Region, and Hemisphere for lateral ROIs and with the Onset Type and Region for midline sites.

As can be seen from Figure 7 (left panel), ERP responses to both onset types were almost identical, displaying the obligatory infant ERP components: P150, N250, and P350 (see Kushnerenko et al., 2002). However, at lateral fronto-central sites, these obligatory components were less pronounced for NP2 than for sentence onset. Accordingly, for the TW 0–500 msec, ANOVAs revealed an Onset Type × Region interaction, $F(2, 66) = 7.19, p \leq .01$, at lateral sites, whereas there was no effect involving onset type in the later TWs 500–1000 msec and 1000–1500 msec. Follow-up one-way ANOVAs yielded main effects of onset type for anterior, $F(2, 66) = 7.08, p \leq .05$, and central regions, $F(2, 66) = 4.75, p \leq .05$. These amplitude differences are

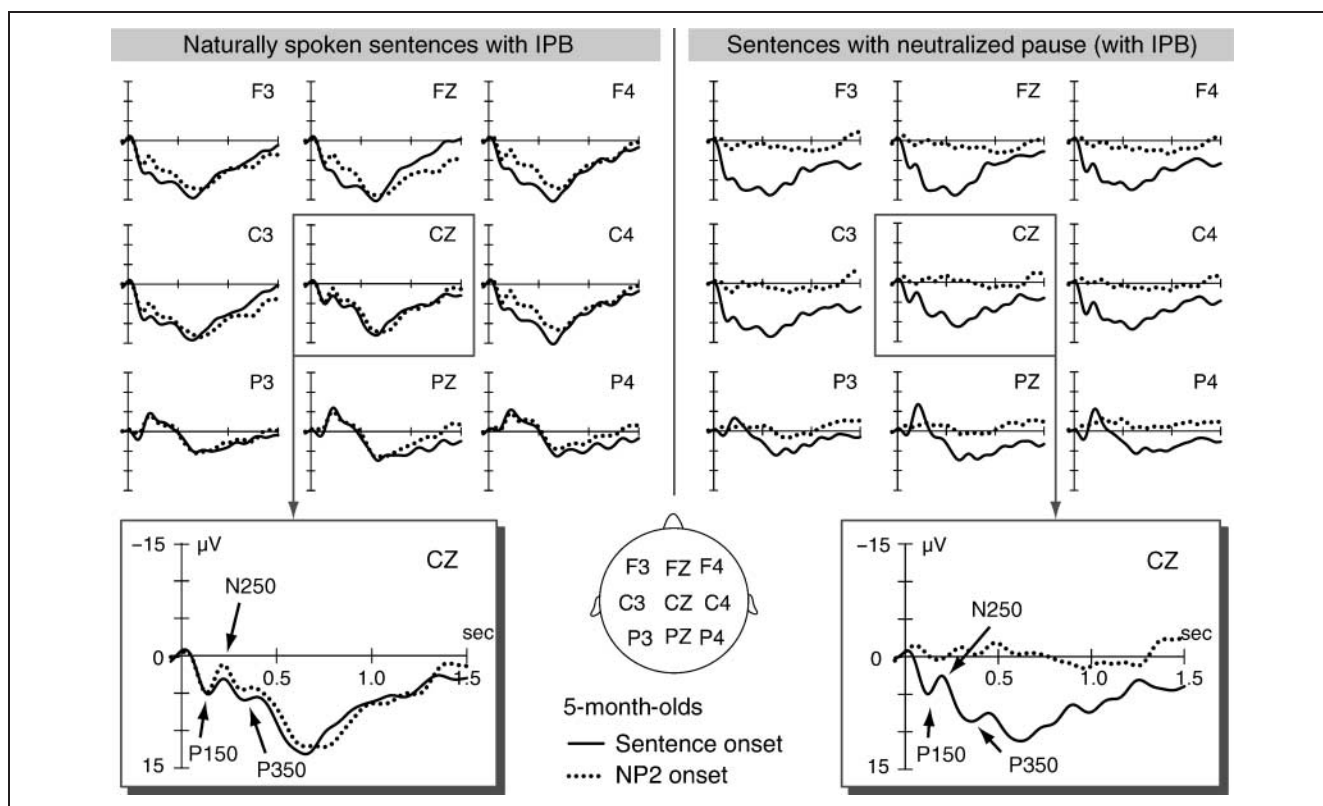


Figure 7. Five-month-olds—sentences with IPB: grand average ERPs to naturally spoken sentences (left) and sentences with neutralized pause (right) relative to sentence onset (solid line) and NP2 onset (dotted line). For naturally spoken sentences, ERPs show obligatory components (P150, N250, and P350) to both onset types; for sentences with neutralized pause, only to the sentence onset.

most likely attributable to differences in stimulus intensity and pause length, as has been shown for the modulation of obligatory components in older children and adults (see Coch, Skenzel, & Neville, 2005; Crowley & Colrain, 2004). Thus, longer pauses (1500 msec preceding sentence onset vs. 560 msec preceding NP2 onsets) and higher intensity (58 dB for sentence onsets vs. 51 dB for NP2 onsets) result in larger ERP amplitudes. A more fine-grained analysis of the onset type effect across TWs of 100 msec revealed that the effect spans the period of 100–400 msec relative to sentence/NP2 onset (Table 4).

In contrast to the naturally spoken sentences with IPB (Experiment 1), the sentences with neutralized pause with IPB (Experiment 2) were found to evoke obligatory ERP responses to sentence onsets but not to NP2 onsets (Figure 7, right panel). This was reflected in a zero-level ERP for NP2 onset processing, whereas the ERP responses relative to sentence onset showed a similar pattern as in Experiment 1. Accordingly, for all three TWs, ANOVAs revealed main effects of onset type at midline sites and lateral ROIs. Additional interaction effects involving onset type mainly demonstrate onset type differences at anterior and central sites (Table 5).

For naturally spoken sentences with IPB, the direct comparison of sentence onset and NP2 onset suggests that the observed positive shift can be accounted for by an obligatory response to the NP2 onset because none of

the expected differences in favor of a CPS were observed. Amplitudes were more pronounced for sentence onset than for NP2 onset, excluding the possibility of additive effects, and furthermore, no amplitude differences occurred at posterior sites where the occurrence of a CPS in response to the IPB, but not obligatory components, would have been expected. However, these results do not yet rule out the possibility of a CPS preceding the obligatory onset response. Therefore, we compared the sentences with neutralized pause (with IPB) and the sentences without IPB relative to NP2 onsets by covering preceding and subsequent TWs. As can be seen for C3, there is no positive shift starting before the IPB (Figure 8). In sum, the positive shift observed in the naturally spoken sentences with IPB is accounted for by the obligatory ERP response to the NP2 onset following the IPB. Importantly, this obligatory response in infants disappears when no boundary pause is present.

The Role of the Pause in Obligatory Onset Responses—Infant and Adult ERP Data

Analyses of infant ERPs to sentences with IPB revealed an obligatory response to NP2 onsets. However, this ERP component only occurred for naturally spoken sentences with pauses at IPBs (Experiment 1), but not for sentences

Table 4. Five-Month-Olds—Naturally Spoken Sentences with IPB: Significant Effects of ANOVAs for the 0- to 500-msec Latency Range Relative to Sentence/NP2 Onset

<i>TW (msec)</i>	<i>Lateral ROIs</i>			<i>Midline Sites</i>		
	<i>Effect</i>	<i>df</i>	<i>F</i>	<i>Effect</i>	<i>df</i>	<i>F</i>
100–200	Onset × Reg	2, 66	5.06*			
200–300	Onset	1, 33	5.43*	Onset × Reg	2, 66	3.44*
	Onset × Reg	2, 66	12.74**	FZ	1, 33	4.15*
	Ant	1, 33	10.57**			
	Cen	1, 33	8.18**			
300–400	Onset	1, 33	4.31*			
	Onset × Reg	2, 66	8.2**			
	Ant	1, 33	9.74**			
	Cen	1, 33	4.75*			
	Onset × Hem	1, 33	5.36*			
	Right	1, 33	7.55**			

Onset = onset type; Reg = region; Hem = hemisphere; Ant = anterior; Cen = central.

* $p \leq .05$.

** $p \leq .01$.

with neutralized pause without pauses at IPBs (Experiment 2). In other words, infants seem to process sentence interruptions caused by pauses, but only when the prosodic break is signaled by a pause. Importantly, the detection of this type of speech interruption reflects low-level acoustic processing rather than the perception of com-

bined prosodic boundary cues to IPs at a cognitive level; the latter would be indicated by a CPS.

From the described results in 5-month-olds, one can derive two possible explanations. The obligatory response is triggered either (a) by the actual pause (and accordingly does not occur when the pause is absent) or (b) by a break

Table 5. Five-Month-Olds—Sentences with Neutralized Pause (with IPB): Significant Effects of ANOVAs for the 0- to 1500-msec Latency Range Relative to Sentence/NP2 Onset

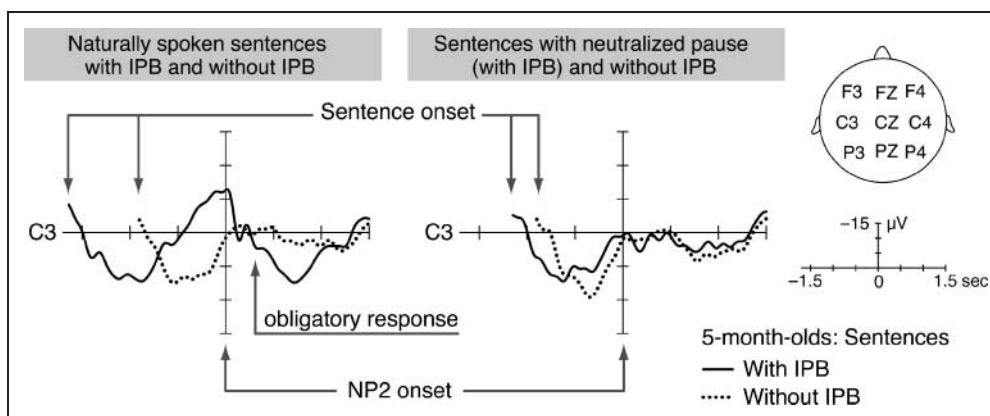
<i>TW (msec)</i>	<i>Lateral ROIs</i>			<i>Midline Sites</i>		
	<i>Effect</i>	<i>df</i>	<i>F</i>	<i>Effect</i>	<i>df</i>	<i>F</i>
0–500	Onset	1, 33	36.07**	Onset	1, 33	16.12**
	Onset × Reg	2, 66	31.69**	Onset × Reg	2, 66	3.44*
	Ant	1, 33	51.47**	FZ	1, 33	22.75**
	Cen	1, 33	43.81**	CZ	1, 33	14.24**
500–1000	Onset	1, 33	17.17**	Onset	1, 33	21.92**
	Onset × Reg	2, 66	6.31**			
	Ant	1, 33	19.64**			
	Cen	1, 33	19.98**			
	Post	1, 33	4.81*			
1000–1500	Onset	1, 33	7.22*	Onset	1, 33	6.0*

Onset = onset type; Reg = region; Ant = anterior; Cen = central; Post = posterior.

* $p \leq .05$.

** $p \leq .01$.

Figure 8. Five-month-olds: grand average ERPs (displayed at C3) to naturally spoken sentences (left) and sentences with neutralized pause (right). Sentences with IPB (solid line) and sentences without IPB (dotted line) are displayed relative to NP2 onsets; sentence onsets are indicated. In sentences with IPB, no positive shift (CPS) is directly preceding the IPB.



signaled by the combined information of several boundary features. When the pause is absent, however, the remaining boundary cues are not sufficient to signal a speech interruption. These assumptions were tested across age groups to tease apart general versus infant-specific processing mechanisms. In adults, both sentence and NP2 onsets of naturally spoken sentences with IPB evoked typical obligatory adult ERP responses, the N1–P2 complex, although they were less pronounced for NP2 onsets (Figure 9, left panel). Interestingly, a similar ERP pattern occurred for

sentence and NP2 onsets of sentences with neutralized pause (with IPB; Figure 9, right panel). Thus, in contrast to infants, the NP2 onsets evoked obligatory ERP components in adults even in the absence of pause at the IPB.

To quantify amplitude and latency differences of the N1 and the P2 components of both onset types across sentences with IPB from both experiments, we extracted peak maxima and minima and their respective latencies in the following TWs: 50–100 msec for N1 and 120–220 msec for P200. Amplitude and latency values are given in Table 6

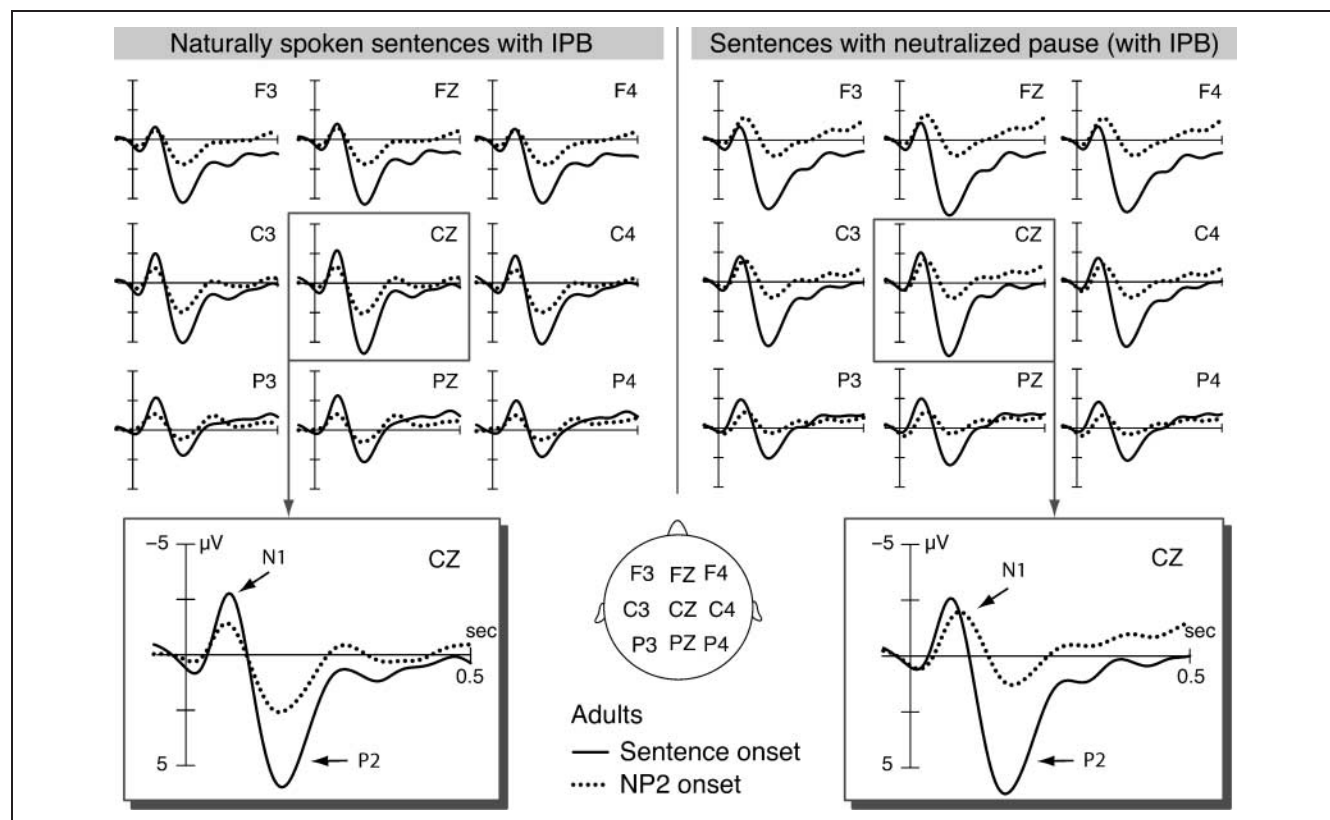


Figure 9. Adults—sentences with IPB: grand average ERPs to naturally spoken sentences (left) and sentences with neutralized pause (right) relative to sentence onset (solid line) and NP2 onset (dotted line). ERPs show obligatory responses (N1–P2 complex) to both onset types for naturally spoken sentences and sentences with neutralized pause.

for naturally spoken sentences with IPB and in Table 7 for sentences with neutralized pause (with IPB).

As the amplitude measures in Tables 6 and 7 indicate, N1 and P2 were generally more pronounced for sentence onsets than for NP2 onsets. These amplitude differences correspond to the effects observed in infants, although with an even broader appearance in adults, and are due to variations in stimulus intensity and pause length (see Coch et al., 2005; Crowley & Colrain, 2004). Accordingly, ANOVAs performed for both experiments revealed main effects and interaction effects involving onset type (see Tables 8 and 9). For naturally spoken sentences with IPB, latency measures provided shorter N1 and P2 latencies for NP2 than for sentence onset responses, whereas for sentences with neutralized pause (with IPB), the reverse pattern applied. As noted, NP2 onsets evoked obligatory components following the IPB in both naturally spoken sentences (Figure 9, left panel) and sentences with neutralized pause (Figure 9, right panel). This was statistically confirmed by multiple one-sample *t* tests against the test value 0, which revealed significant deviations for all amplitude measures at midline and lateral sites (Tables 6 and 7, column 5).

Figure 10 (left panel) convincingly illustrates the similarity of the NP2 onset responses to the naturally spoken sentences with IPB and the sentences with neutralized pause (with IPB) for adult listeners. In contrast, the infant ERP data demonstrate the described discrepancy between ERP responses to NP2 onsets in the presence and the absence of the boundary pause (Figure 10, right panel). For statistical comparisons of the adult NP2 onset responses across experiments, we performed ANOVAs with the additional between-subject factor, Experiment, for N1 and P2 amplitude and latency measures at lateral and midline sites. ANOVAs revealed Experiment \times Region interaction effects for N1 amplitudes at lateral, $F(2, 106) = 6.45, p \leq .01$, and midline sites, $F(2, 106) = 7.79, p \leq .01$. Subsequent independent two-sample *t* tests showed that these differences were restricted to anterior lateral ROIs, $t(53) = 2.42, p \leq .05$, and FZ, $t(53) = 2.39, p \leq .05$. The similarity of the NP2 onset responses across experiments was further supported by ANOVAs calculated for N1–P2 peak-to-peak measures. Here, no peak-to-peak latency differences were obtained across experiments. Peak-to-peak amplitude differences were reflected in an Experiment \times Region interaction at midline electrodes,

Table 6. Adults—Naturally Spoken Sentences with IPB: N1 and P2 Peak Amplitudes and Latencies (*SD* Indicated)

<i>TW</i>	<i>ROIs/Sites</i>	<i>Sentence Onset</i>		<i>NP2 Onset</i>	
		<i>Amp</i> (μ V)	<i>Lat</i> (<i>msec</i>)	<i>Amp</i> (μ V)	<i>Lat</i> (<i>msec</i>)
N1	Right ant	−1.40 (2.18)	77 (9)	−1.28 (1.27)*	74 (12)
	Left ant	−1.58 (2.24)	76 (10)	−1.43 (1.03)*	74 (13)
	Right cen	−2.00 (1.84)	80 (8)	−1.38 (0.95)*	78 (12)
	Left cen	−2.33 (2.11)	78 (9)	−1.68 (1.06)*	74 (12)
	Right post	−2.71 (1.73)	82 (12)	−1.37 (0.98)*	78 (13)
	Left post	−3.06 (1.76)	82 (13)	−1.63 (1.16)*	75 (13)
	FZ	−2.18 (2.80)	77 (12)	−1.75 (1.38)*	75 (14)
	CZ	−3.68 (3.22)	77 (9)	−2.24 (1.41)*	73 (14)
	PZ	−3.77 (2.42)	80 (11)	−2.11 (1.36)*	73 (15)
P2	Right ant	5.78 (2.81)	172 (16)	2.93 (2.28)*	174 (20)
	Left ant	5.47 (2.63)	171 (16)	2.72 (2.13)*	172 (21)
	Right cen	4.88 (2.01)	174 (16)	2.85 (1.69)*	170 (21)
	Left cen	4.53 (2.31)	180 (16)	2.32 (1.68)*	175 (17)
	Right post	1.90 (1.54)	176 (18)	1.33 (1.42)*	168 (21)
	Left post	1.72 (1.78)	180 (19)	0.96 (1.55)*	169 (18)
	FZ	6.13 (3.36)	173 (17)	2.90 (2.64)*	168 (25)
	CZ	6.73 (3.79)	169 (16)	3.51 (2.83)*	166 (25)
	PZ	3.55 (2.79)	175 (22)	1.90 (2.27)*	170 (23)

Amp = amplitude; *Lat* = latency; ant = anterior; cen = central; post = posterior.

* $p \leq .01$; one-sample *t* test against the test value 0.

Table 7. Adults—Sentences with Neutralized Pause (with IPB): N1 and P2 Peak Amplitudes and Latencies (*SD* Indicated)

<i>TW</i>	<i>ROIs/Sites</i>	<i>Sentence Onset</i>		<i>NP2 Onset</i>	
		<i>Amp</i> (μV)	<i>Lat</i> (<i>msec</i>)	<i>Amp</i> (μV)	<i>Lat</i> (<i>msec</i>)
N1	Right ant	−1.61 (1.84)	72 (10)	−2.07 (1.39)*	82 (12)
	Left ant	−1.64 (1.83)	72 (8)	−2.22 (1.31)*	82 (12)
	Right cen	−1.93 (1.71)	75 (11)	−1.53 (1.31)*	80 (13)
	Left cen	−2.35 (1.51)	74 (12)	−1.82 (1.28)*	83 (11)
	Right post	−2.48 (1.42)	76 (12)	−1.40 (1.19)*	81 (14)
	Left post	−2.69 (1.41)	77 (11)	−1.60 (1.07)*	84 (15)
	FZ	−2.20 (2.19)	73 (9)	−2.72 (1.62)*	83 (16)
	CZ	−3.39 (2.28)	74 (9)	−2.67 (1.75)*	84 (15)
	PZ	−3.16 (1.85)	75 (13)	−1.98 (1.48)*	83 (15)
P2	Right ant	6.44 (2.36)	174 (17)	1.88 (1.34)*	186 (18)
	Left ant	5.99 (2.33)	175 (20)	1.85 (1.62)*	181 (17)
	Right cen	5.21 (1.79)	176 (15)	1.89 (1.19)*	181 (18)
	Left cen	4.74 (2.37)	177 (12)	1.72 (1.39)*	181 (17)
	Right post	2.29 (1.89)	176 (17)	1.12 (1.41)*	173 (18)
	Left post	2.01 (2.31)	175 (17)	1.03 (1.47)*	172 (20)
	FZ	7.00 (3.10)	175 (20)	1.87 (1.95)*	185 (21)
	CZ	6.77 (3.46)	169 (17)	2.02 (1.86)*	178 (20)
	PZ	3.76 (2.6)	172 (20)	1.22 (1.88)*	172 (22)

Amp = amplitude; *Lat* = latency; ant = anterior; cen = central; post = posterior.

* $p \leq .01$; one-sample *t* test against the test value 0.

Table 8. Adults—Naturally Spoken Sentences with IPB: Significant Effects of ANOVAs for N1 and P2 Peak Amplitudes and Latencies

<i>TW</i>	<i>Lateral ROIs</i>			<i>Midline Sites</i>		
	<i>Effect</i>	<i>df</i>	<i>F</i>	<i>Effect</i>	<i>df</i>	<i>F</i>
N1-Amp	Onset	1, 26	7.57*	Onset	1, 26	7.68**
	Onset \times Reg	2, 52	14.56**	Onset \times Reg	2, 52	9.31**
	Cen	1, 26	5.2*	CZ	1, 26	7.37*
	Post	1, 26	28.17**	PZ	1, 26	18.21**
N1-Lat	Onset	1, 26	5.03*			
P2-Amp	Onset	1, 26	64.59**	Onset	1, 26	60.33**
	Onset \times Reg	2, 52	35.55**	Onset \times Reg	2, 52	12.91**
	Ant	1, 26	133.11**	FZ	1, 26	106.23**
	Cen	1, 26	58.16**	CZ	1, 26	56.48**
	Post	1, 26	4.86*	PZ	1, 26	13.07**
P2-Lat	Onset \times Reg	2, 52	4.44*			
	Post	1, 26	6.52*			

Amp = amplitude; *Lat* = latency; *Onset* = onset type; *Reg* = region; *Ant* = anterior; *Cen* = central; *Post* = posterior.

* $p \leq .05$.

** $p \leq .01$.

Table 9. Adults—Sentences with Neutralized Pause (with IPB): Significant Effects of ANOVAs for N1 and P2 Peak Amplitudes and Latencies

TW	Lateral ROIs			Midline Sites		
	Effect	df	F	Effect	df	F
N1-Amp	Onset × Reg	2, 54	26.77**	Onset × Reg	2, 54	22.85**
	Post	1, 27	13.37**	PZ	1, 27	8.78**
N1-Lat	Onset	1, 27	9.17**	Onset	1, 27	7.26*
P2-Amp	Onset	1, 27	63.64**	Onset	1, 27	62.33**
	Onset × Reg	2, 54	74.13**	Onset × Reg	2, 54	31.02**
	Ant	1, 27	130.06**	FZ	1, 27	103.66**
	Cen	1, 27	71.93**	CZ	1, 27	57.72**
	Post	1, 27	6.62*	PZ	1, 27	20.98**
P2-Lat	Onset × Reg	2, 54	6.52**	Onset × Reg	2, 54	3.31*
	Ant	1, 27	5.11*			

Amp = amplitude; Lat = latency; Onset = onset type; Reg = region; Ant = anterior; Cen = central; Post = posterior.

* $p \leq .05$.

** $p \leq .01$.

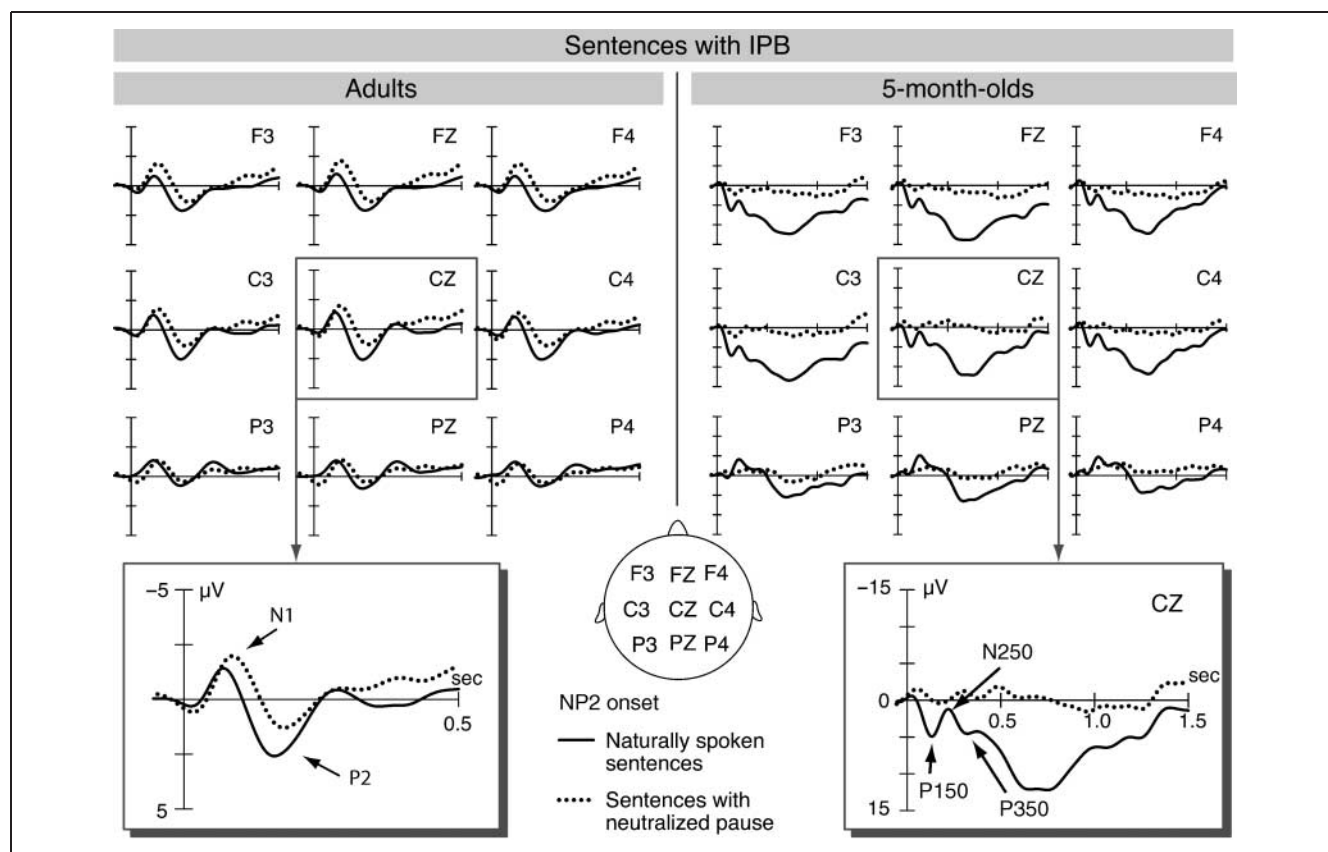


Figure 10. NP2 onset: grand average ERPs of adults (left panel) and 5-month-olds (right panel) to NP2 onsets of naturally spoken sentences (solid line) and sentences with neutralized pause (dotted line). In adults, ERPs display obligatory components to both sentence types; in infants, ERPs show obligatory component only to naturally spoken sentences.

$F(2, 106) = 4.17, p \leq .05$, for which, however, subsequent independent two-sample t tests delivered no significant effects at single electrode sites. In summary, from the study of adult ERP responses using the same sentence material as for infants, we can conclude that adults' perceptions of speech interruptions are not solely based on pausing but are the result of a combination of prosodic boundary features. Thus, even when the pause at the IPB is deleted, adults still show obligatory onset responses, with pre-boundary lengthening and pitch change being sufficient cues for signaling prosodic breaks.

DISCUSSION

The present study investigated the neurophysiology underlying how 5-month-old German-learning infants process prosodic boundaries by measuring the associated ERPs. In addition, we probed for developmental differences between infants and adults by presenting both cohorts with sentences that were either naturally spoken or that had no pause cue at the IPB.

For adults, we observed CPS and obligatory N1–P2 responses to both naturally spoken sentences with IPB and sentences with neutralized pause (with IPB). Naturally spoken sentences with IPB contained all of the available prosodic boundary information, that is, pitch change, pre-boundary lengthening, and pausing, whereas sentences with neutralized pause (with IPB) retained the effects due to pitch change and preboundary lengthening. Thus, IPBs evoked boundary perception (and the lower level detection of an acoustic interruption) even when the pause at the IPB was deleted. In other words, preboundary lengthening and pitch change are sufficient to signal prosodic breaks, independent of the presence of a boundary pause.⁶ These results are in agreement with previously reported CPS in IP processing in adults (see Steinhauer, 2003; Steinhauer et al., 1999).

In infants, the ERP patterns to the naturally spoken sentences with IPB also exhibited a positive shift in response to IPBs. This positive shift, however, did not reflect an adult-like CPS but instead was an infant obligatory ERP response to the onset of speech after the IPB. We surmised that infants detect speech interruptions by low-level acoustic processes rather than by higher order perception of combined prosodic boundary cues to IPs. In contrast to the naturally spoken sentences, the sentences with neutralized pause (with IPB) did not elicit a positive shift in the ERP. Thus, the obligatory response to IPBs disappeared when no boundary pause was present, indicating that infants process sentence interruptions caused by acoustic cues only when the prosodic break is signaled by a pause.

There are two possible explanations for our observations regarding the processing of prosodic boundaries in young infants. First, German-learning 5-month-olds do not yet process IPBs. Instead, infants' detection of speech in-

terruptions is driven by the pause and, consequently, does not occur when the pause is absent. Alternatively, German-learning infants at this age may identify prosodic breaks composed of combined boundary cues but only do so if speech interruptions are signaled by a pause. In contrast to adults, these processes happen at a lower level perceptual stage and, therefore, do not elicit a CPS. Independent of which explanation is correct, our results offer two conclusions. First, infants' processing of breaks in the speech signal happens at a lower order perceptual level, and second, the pause is a particularly salient cue in the processing of those breaks by German infants.

Regarding the first interpretation, our results are in disagreement with behavioral data that have previously described how English-learning infants process prosodic boundary information. These studies can be roughly grouped into two lines of research. First, studies using the pause insertion technique have demonstrated that infants prefer sequences containing pauses at clause/phrase boundary positions over passages with pauses inserted elsewhere in the utterance (Juszyk et al., 1992, 1995; Hirsh-Pasek et al., 1987). The second line of research has described how infants prefer familiar sequences presented in larger test passages when these sequences occurred as syntactic units (within prosodic boundaries), as opposed to when they occurred in syntactic nonunits (across prosodic boundaries; Seidl, 2007; Soderstrom et al., 2003, 2005; Nazzi et al., 2000). For the latter studies, the authors concluded that infants do not only perceive prosodic cues but actually use them for the detection of units in continuous speech.

Given the assumption that our ERP study and the described behavioral studies test for similar prosodic processing abilities, cross-linguistic differences seem to be a critical factor because German 5-month-olds, in contrast to English 6-month-olds, seem unable to process prosodic boundaries. The interpretation of cross-linguistic differences might be best approached by the evaluation of differences in the German and the English intonation systems (see Gibbon, 1998). The functional demands on prosody are likely to be lower in German than in English (see Schubiger, 1980) because German has a larger number of inflections and a relatively flexible word order. Word order serves as focalization device because it is applied in topicalizations and used to denote new information. Furthermore, German uses a large number of discourse particles that probably take over functions represented by intonation patterns in other languages (see Schubiger, 1980). Consequently, German features a lower number of intonation patterns and a reduced range of pitch variation, relative to English, which results in less variation across boundary tones (Markus, 2006; Gibbon, 1998). Indeed, in a cross-linguistic analysis of prosodic cues in infant-directed speech, Fernald et al. (1989) found that American English parents used the most exaggerated speech (international modifications) as compared with, for example, German-, French-, and Italian-speaking parents. Thus, it might be the case that German and English infants differ

in their ability to process prosodic breaks based on the respective strength of IP marking in their native language.

The second interpretation of our data is that German-learning infants process prosodic breaks via lower level perceptual processes, but only when the speech interruption is signaled by a pause. Further studies comparing pause perception at prosodic boundary positions and nonboundary positions will be needed to determine the particular role of pausing in infants' IPB processing. However, the current results already point to the generally important role of pausing in the detection of speech interruptions. A study by Seidl (2007) revealed that English-learning infants were able to perceive IPs when presented with sentences lacking pause as boundary information. The fact that the absence of the pause cue seems to matter to German-learning but not to English-learning infants again points to cross-linguistic differences, particularly in the weighting of acoustic cues.⁷ As mentioned earlier, the flexible German word order takes over functions that are accomplished by particular pitch patterns in other languages. Thus, it seems plausible that pausing in German plays a role as an additional syntactic structuring device (see also Butcher, 1981). Furthermore, less intonation patterns and less pitch variations are likely to yield less modulated boundary tones, which without being marked by pauses, may not be sufficient for triggering boundary detection.

The comparison of the infant and the adult ERP data suggests two types of developmental differences: first, differences in the underlying neurophysiological processes, and second, differences in the role of the pause as an acoustic boundary cue. Concerning the latter, the differential weighting of acoustic cues might change as language acquisition progresses. Children who have successfully identified basic syntactic structure are less dependent on prosodic markers, so pausing may no longer be as important for them. However, reliance on pause as a boundary cue can still be observed in adults when the language input is not well phrased and is prosodically ambiguous (see Dankovicová, Pigott, Wells, & Peppé, 2004). Additional evidence for the significant role of pauses in language learning comes from studies on the acquisition of artificial grammar. For instance, syntax acquisition in German-speaking adults was modeled by the presentation of rule-based versus random strings of pseudowords, which demonstrated how rule acquisition is bolstered by the presence of pauses placed between the units of individual sequences (Mueller, Bahlmann, & Friederici, 2008). Although rule learning can also occur without pauses, it is more effective when pauses are present (see also Peña, Bonatti, Nespor, & Mehler, 2002).

Regarding developmental differences in the underlying neurophysiological processes, our ERP data suggest that German infants primarily incorporate prosodic boundary information by using low-level sensory processing.⁸ In accordance with prosodic bootstrapping, infants initially detect salient acoustic cues in the speech input like silences

or large pitch changes. These cues signal where constituents are and eventually enable infants to bootstrap into other cues associated with the edges of these constituents. Thus, as language acquisition progresses, children learn that certain acoustic cues correlate and conjointly mark structural units. We hypothesize that these processes are characterized by different underlying neurophysiological mechanisms. For instance, attention to salient acoustic cues is characterized by sensory processing mechanisms (obligatory ERP responses), whereas perception of a boundary as a correlation of several cues (i.e., one integrated percept) is characterized by cognitive processing mechanisms (CPS response). Thus, the acquisition of the concept of a prosodic boundary is a gradual process that develops out of the early response to acoustically salient cues, such as pause. Further developmental studies are needed to determine when the processes indicated by the CPS emerge⁹ and under which circumstances they are likely to be elicited, for example, boundary strength and presence of certain boundary cues.

In sum, the existing data suggest that during early development, German-learning infants are able to detect particularly salient acoustic cues in the speech input, which are likely contributors to the recognition of prosodic boundaries. Later, these early processing tendencies develop so that infants can recognize prosodic boundaries based on combined acoustic information and ultimately use this knowledge to extract structural units from continuous speech.

Conclusion

In conclusion, the current study provides neurophysiological evidence for 5-month-old German-learning infants to process sentence internal pauses, as revealed by obligatory ERP responses following the IPBs. The fact that prosodic breaks elicit obligatory ERP components in infants, but not an adult-like CPS, suggests a strong reliance on low-level sensory processing during early infancy. Furthermore, the current data of German-learning infants point to developmental and cross-linguistic differences in the weighting of particular boundary cues in IP processing.

Acknowledgments

We would like to thank Christina Rügen and Jördis Haselow for their dedication and patience in working with infants and their laborious EEG data collection. Furthermore, we are thankful to the participating families for their persistent commitment to our research. We are grateful to Dr. Volker Hesse and Dr. Sebastian Buttenberg, heads of the neonatology department of the gynecological hospital (Sana Klinikum Berlin Lichtenberg), for making it possible to contact parents directly. We are indebted to an anonymous reviewer, who identified crucial additional analyses. Finally, we would like to thank Dr. Johanna Barry and Scott Greenwald for their valuable comments and improvements upon the first draft of the manuscript. The current research was supported, in part, by the EU (EC12778/NEST-CALACEI Project).

Reprint requests should be sent to Claudia Männel, Max Planck Institute for Human Cognitive and Brain Sciences, PO Box 500 355, 04303 Leipzig, Germany, or via e-mail: maennel@cbs.mpg.de.

Notes

1. Note that due to the indirect relationship between syntax and prosody, there is only a rough correspondence between prosodic units (intonational and phonological phrases) and syntactic units (clauses and noun/verb phrases).
2. These ERP components are called “obligatory” components because they are invariably elicited when input reaches the sensory systems. Accordingly, these components are modulated by the physical properties of the eliciting stimuli.
3. Note that this is different for the English version of the example sentences. Here, the word order directly indicates the syntactic role of *Papa: Tommi promises to help papa* versus *Tommi promises papa to sleep*. Thus, the prosodic realization of English sentences does not play the same disambiguating role as it does in German sentences.
4. The condition effect in the latter TW is certainly amplified by the contrarily proceeding ERPs to sentences without IPB. For this TW, the ERP to this generally shorter sentence type most likely merges into a pronounced positivity (starting with a negative deflection) in response to the sentence end prosody.
5. This relatively short prestimulus baseline was chosen because this was about the minimal pause length (for both Experiments 1 and 2) preceding Part 2. Thus, only for this short period were there no condition differences between the prestimulus intervals of the two onset types.
6. Note that the fact that in adults pauses are not necessary for IPB perception in the presence of other boundary cues does not exclude the notion that pausing becomes relevant in the absence of other cues (see cue trading; Beach, 1991).
7. The possibility that German-learning infants would nonetheless be able to perform phrase segmentation at the behavioral level seems unlikely given that they did not show an obligatory effect to the IPB at the neurophysiological level. During the developmental course, neurophysiological evidence typically precedes behavioral performance and not vice versa.
8. For a related discussion of perceptual mechanisms in language learning, see Mehler, Endress, Gervain, and Nespor (2008). The authors illustrate how the findings of most studies on AGL can be explained based on so-called *perceptual primitives*, such that language learners direct their attention to repetitions and acoustically marked edges of speech units in the language input.
9. We refrain from the discussion of a developmental shift between the age of 5 and 8 months, under the assumption that similar low-level perceptual mechanisms are active in 8-month-olds (Pannekamp, Weber, & Friederici, 2006).

REFERENCES

Aasland, W. A., & Baum, S. R. (2003). Temporal parameters as cues to phrasal boundaries: A comparison of processing by left- and right-hemisphere brain-damaged individuals. *Brain and Language*, *87*, 385–399.

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., & Edwards, J. (1990). Lengthening and shortening and the nature of prosodic constituency. In J. Kingston & M. Beckman (Eds.), *Papers in laboratory phonology*

I. Between the grammar and the physics of speech (pp. 152–178). Cambridge: Cambridge University Press.

Butcher, A. (1981). *Aspects of the speech pause: Phonetic correlates and communicative functions* (Arbeitsberichte Nr. 15). Kiel: Institut für Phonetik und digitale Sprachverarbeitung, Universität Kiel.

Chatrian, G. E., Lettich, E., & Nelson, P. L. (1988). Modified nomenclature for the “10%” electrode system. *Journal of Clinical Neurophysiology*, *5*, 183–186.

Christophe, A., Dupoux, E., Bertoncini, J., & Mehler, J. (1994). Do infants perceive word boundaries? An empirical study of the bootstrapping of lexical acquisition. *Journal of the Acoustical Society of America*, *95*, 1570–1580.

Christophe, A., Mehler, J., & Sebastián-Gallés, N. (2001). Perception of prosodic boundary correlates by newborn infants. *Infancy*, *2*, 385–394.

Coch, D., Skenzel, W., & Neville, H. J. (2005). Auditory and visual refractory periods effects in children and adults: An ERP study. *Clinical Neurophysiology*, *116*, 2184–2203.

Cooper, R., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development*, *61*, 1584–1595.

Cooper, W., & Paccia-Cooper, J. (1980). *Syntax and speech*. Cambridge, MA: Harvard University Press.

Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: Age, sleep and modality. *Clinical Neurophysiology*, *115*, 732–744.

Dankovicová, J., Pigott, K., Wells, B., & Peppé, S. (2004). Temporal markers of prosodic boundaries in children’s speech production. *Journal of the International Phonetic Association*, *34*, 17–36.

Delattre, P. (1966). A comparison of syllable length conditions across languages. *International Review of Applied Linguistics*, *IV*, 183–198.

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior & Development*, *8*, 181–195.

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., De Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers’ and fathers’ speech to preverbal infants. *Journal of Child Language*, *16*, 477–501.

Fisher, C., & Tokura, H. (1996). Acoustic cues to grammatical structure in infant-directed speech: Cross-linguistic evidence. *Child Development*, *67*, 3192–3218.

Friederici, A. D., & Thierry, G. (2008). *Early language development: Bridging brain and behaviour. Trends in language acquisition research (TiLAR)* (Vol. 5). Amsterdam: John Benjamins.

Gerken, L. (1996). Phonological and distributional cues to syntax acquisition. In J. Morgan & K. Demuth (Eds.), *Signal to syntax* (pp. 411–425). Mahwah, NJ: Erlbaum.

Gerken, L., Jusczyk, P. W., & Mandel, D. (1994). When prosody fails to cue syntactic structure: 9-month-olds’ sensitivity to phonological versus syntactic phrases. *Cognition*, *51*, 237–265.

Gibbon, D. (1998). Intonation in German. In D. Hirst & A. Di Cristo (Eds.), *Intonation systems: A survey of twenty languages* (pp. 78–95). Cambridge: Cambridge University Press.

Gleitman, L., Gleitman, H., Landau, B., & Wanner, E. (1988). Where the learning begins: Initial representations for language learning. In F. Newmeyer (Ed.), *The Cambridge Linguistic Survey* (pp. 150–193). Cambridge: Cambridge University Press.

Gleitman, L., & Wanner, E. (1982). The state of the state of the art. In E. Wanner & L. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 3–48). Cambridge, MA: Cambridge University Press.

Gout, A., Christophe, A., & Morgan, J. L. (2004). Phonological phrase boundaries constrain lexical access: II. Infant data. *Journal of Memory and Language*, *51*, 548–567.

- Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*, 95–112.
- Hirsh-Pasek, K., Kemler Nelson, D. G., Jusczyk, P. W., Wright Cassidy, K., Druss, B., & Kennedy, L. (1987). Clauses are perceptual units for young infants. *Cognition*, *26*, 269–286.
- Johnson, E. K., & Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language*, *44*, 1–20.
- Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, *29*, 1–23.
- Jusczyk, P. W., Hirsh-Pasek, K., Nelson, D. G., Kennedy, L. J., Woodward, A., & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology*, *24*, 252–293.
- Jusczyk, P. W., Hohne, E., & Mandel, D. (1995). Picking up regularities in the sound structure of the native language. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research* (pp. 91–119). Timonium, MD: York Press.
- Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, *39*, 159–207.
- Kerkhofs, R., Vonk, W., Schriefers, H., & Chwilla, D. J. (2007). Discourse, syntax, and prosody: The brain reveals an immediate interaction. *Journal of Cognitive Neuroscience*, *19*, 1421–1434.
- Kushnerenko, E., Ceponiene, R., Balan, P., Fellman, V., Huutilainen, M., & Näätänen, R. (2002). Maturation of the auditory event-related potentials during the first years of life. *NeuroReport*, *13*, 47–51.
- Ladd, D. R. (1988). Declination “reset” and the hierarchical organization of utterances. *Journal of the Acoustical Society of America*, *84*, 530–554.
- Markus, M. (2006). English and German prosody: A contrastive comparison. In Y. Kawaguchi, I. Fonagy, & T. Moriguchi (Eds.), *Prosody and syntax: Cross-linguistic perspectives* (pp. 103–124). Amsterdam: John Benjamins.
- Mehler, J., Endress, A., Gervain, J., & Nespors, M. (2008). From perception to grammar. In A. D. Friederici & G. Thierry (Eds.), *Early language development: Bridging brain and behaviour. Trends in language acquisition research (TiLAR)* (Vol. 5, pp. 191–213). Amsterdam: John Benjamins.
- Mehler, J., Jusczyk, P. W., Lambert, G., Halsted, G., Bertoincini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, *29*, 143–178.
- Morgan, J. L., Meier, R. P., & Newport, E. L. (1987). Structural packaging in the input to language learning: Contributions of prosodic and morphological marking of phrases to the acquisition of language. *Cognitive Psychology*, *19*, 498–550.
- Morgan, J. L., & Saffran, J. R. (1995). Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. *Child Development*, *66*, 911–936.
- Mueller, J. L., Bahlmann, J., & Friederici, A. D. (2008). The role of pause cues in language learning: The emergence of ERPs related to sequence processing. *Journal of Cognitive Neuroscience*, *20*, 892–905.
- Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology*, *24*, 375–425.
- Nazzi, T., Dilley, L. C., Jusczyk, A. M., Shattuck-Hufnagel, S., & Jusczyk, P. W. (2005). English-learning infants' segmentation of verbs from fluent speech. *Language and Speech*, *48*, 279–298.
- Nazzi, T., Kemler Nelson, D. G., Jusczyk, P. W., & Jusczyk, A. M. (2000). Six-month-olds' detection of clauses embedded in continuous speech: Effects of prosodic well-formedness. *Infancy*, *1*, 123–147.
- Nespor, M., & Vogel, I. (1986). *Prosodic phonology*. Dordrecht: Foris.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia*, *9*, 97–113.
- Pannekamp, A., Toepel, U., Alter, K., Hahne, A., & Friederici, A. D. (2005). Prosody-driven sentence processing: An event-related brain potential study. *Journal of Cognitive Neuroscience*, *17*, 407–421.
- Pannekamp, A., Weber, C., & Friederici, A. D. (2006). Prosodic processing at the sentence level in infants. *NeuroReport*, *17*, 675–678.
- Peña, M., Bonatti, L. L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science*, *298*, 604–607.
- Peters, A. (1983). *The units of language acquisition*. Cambridge: Cambridge University Press.
- Pierrehumbert, J. (1980). *The phonology and phonetics of English intonation*. MIT Linguistics PhD thesis, Indiana University Linguistics Club, Bloomington, IN.
- Saffran, J., Newport, E. L., & Aslin, R. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, *35*, 606–621.
- Schubiger, M. (1980). English intonation and German modal particles: II. A comparative study. In L. R. Waugh & C. H. van Schooneveld (Eds.), *The melody of language* (pp. 279–298). Baltimore, MD: University Park Press.
- Scott, D. (1982). Duration as a cue to the perception of a phrase boundary. *Journal of the Acoustical Society of America*, *71*, 996–1007.
- Seidl, A. (2007). Infants' use and weighting of prosodic cues in clause segmentation. *Journal of Memory and Language*, *57*, 24–48.
- Seidl, A., & Johnson, E. (2007). Boundary alignment facilitates 11-month-olds' segmentation of vowel-initial words from speech. *Journal of Child Language*, *34*, 1–24.
- Selkirk, E. (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.
- Soderstrom, M., Nelson, D. G. K., & Jusczyk, P. W. (2005). Six-month-olds recognize clauses embedded in different passages of fluent speech. *Infant Behavior & Development*, *28*, 87–94.
- Soderstrom, M., Seidl, A., Nelson, D. G. K., & Jusczyk, P. W. (2003). The prosodic bootstrapping of phrases: Evidence from prelinguistic infants. *Journal of Memory and Language*, *49*, 249–267.
- Steinhauer, K. (2003). Electrophysiological correlates of prosody and punctuation. *Brain & Language*, *86*, 142–164.
- Steinhauer, K., Alter, K., & Friederici, A. D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience*, *2*, 191–196.
- Streeter, L. A. (1978). Acoustic determinants of phrase boundary location. *Journal of the Acoustical Society of America*, *64*, 1582–1592.
- Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America*, *91*, 1707–1717.