Prosodic phrase boundary perception in adults and infants

What the brain reveals about contextual influence
and the impact of prosodic cues

Kumulative Dissertationsschrift
zur Erlangung des akademischen Grades Doctor philosophiae (Dr. phil.)
im Fach Kognitionswissenschaft

eingereicht bei der
Humanwissenschaftlichen Fakultät
der Universität Potsdam
im Februar 2017

vorgelegt von
Julia Holzgrefe-Lang

Supervision: Prof. Dr. Isabell Wartenburger

1st reviewer: Prof. Dr. Isabell Wartenburger
2nd reviewer: Prof. Dr. Mathias Scharinger

Day of submission: February 3, 2017
Thesis defense: Potsdam, June 28, 2017

Published online at the
Institutional Repository of the University of Potsdam:
URN urn:nbn:de:kobv:517-opus4-405943
http://nbn-resolving.de/urn:nbn:de:kobv:517-opus4-405943

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Ich möchte all denjenigen herzlich danken, die mit ihrer wertvollen Unterstützung zum Gelingen dieser Arbeit beigetragen haben. Insbesondere danke ich Isabell Wartenburger, die meine Dissertation betreut hat, für die kontinuierliche Begleitung, für ihren konstruktiv-kritischen Blick, ihre inhaltlichen Impulse und nicht zuletzt ihr offenes Ohr.

Ein Großteil dieser Arbeit erfolgte im Rahmen eines DFG-Projekts zur prosodischen Kompetenz von Säuglingen und Erwachsenen. In diesem Projekt und darüber hinaus habe ich eng mit Caroline Wellmann zusammengearbeitet, der ich dafür, für den fachlichen Austausch und die entstandene Freundschaft von Herzen danken möchte. Barbara Höhle hat sich als Co-Projektleiterin über die Jahre hinweg immer wieder gewinnbringend mit meiner Arbeit auseinandergesetzt und war nicht nur während Isabells Elternzeit stets eine kompetente und interessierte Ansprechpartnerin für mich. Dafür sei ihr herzlich gedankt! Ich danke weiterhin Hubert Truckenbrodt und Caterina Petrone für die Zusammenarbeit im Projekt sowie für ihre Hilfe bei phonetischen und phonologischen Fragen. Caterina Petrone danke ich außerdem für die Unterstützung bei der statistischen Auswertung der Verhaltensdaten.

Auch Reinhold Kliegl möchte ich herzlich für seine Statistik-Beratung danken, ebenso Ruben van de Vijver für seine Starthilfe bei meinem Einstieg in den wissenschaftlichen Schreibprozess. Gerrit Kentner ist nicht nur ein langjähriger Kommilitone und Freund, sondern auch mein Telefon-Joker für den Bereich Phonologie. Vielen Dank dafür! Als Betreuer meiner Diplomarbeit hat Heiner Drenhaus mein Interesse an der neurokognitiven Sprachverarbeitung gefördert und mich mit der Erhebung ereigniskorrelierter Potentiale vertraut gemacht. Dafür möchte ich mich an dieser Stelle nochmals herzlich bedanken. Regine Oberecker und insbesondere Christina Rügen danke ich für die wertvollen Erfahrungen, die ich während eines Praktikums im Labor für Kindersprachforschung im Kinderkrankenhaus Lindenhof in Berlin sammeln konnte.

Ich hatte das Glück, in Potsdam in einer Arbeitsgruppe zu promovieren, in der die Zusammenarbeit von einem freundschaftlichen Klima, Interesse am fachlichen Austausch sowie gegenseitiger Unterstützung und Motivation geprägt war. Dafür möchte ich allen Mitgliedern der AG Neurokognition der Sprache bzw. Kognitive Neurolinguistik/ Patholinguistik sehr herzlich danken. Insbesondere danke ich meinen Co-Doktorandinnen Juliane Burmester, Romy Räling, Sandra Hanne und

Judith Heide. Dieses Arbeitsumfeld habe ich nach meinem Umzug nach Leipzig sehr vermisst. Umso schöner war es, bei jedem meiner zahlreichen Besuche in Golm so herzlich willkommen zu sein. Ganz besonders danke ich in diesem Zusammenhang Judith für meinen Potsdamer Zweitwohnsitz!

Bei Romy Räling und Sandra Hanne möchte ich mich herzlich dafür bedanken, dass ich von ihren Erfahrungen profitieren durfte und mit guten Tipps für den Endspurt versorgt wurde. Ein besonders großer Dank geht an Hanni für das geduldige, kritische und konstruktive Korrekturlesen meiner Synopsis.

Laura Hahn, Luisa Scherf, Susanne Pelke und Christiane Patzwald danke ich für ihr Interesse an meiner Arbeit, die sie im Rahmen ihrer Abschlussarbeiten und als studentische Hilfskräfte unterstützt haben. Außerdem danke ich allen Studierenden, die an den Experimenten teilgenommen haben, und ganz besonders den jungen Familien, die durch ihre bereitwillige Teilnahme meine Forschung zur frühkindlichen Sprachperzeption erst ermöglicht haben.

Die vorliegende Arbeit wurde durch ein sechsmonatiges Promotionsstipendium der Graduiertenförderung des Landes Brandenburg finanziell unterstützt. Dafür danke ich der Vergabekommission an der Universität Potsdam sehr herzlich.

Zu meiner großen Freude habe ich mittlerweile an der Universität Leipzig eine neue berufliche Heimat gefunden. Es wurde also höchste Zeit, auch das hier vorliegende Kapitel meiner Potsdamer Zeit abzuschließen. Ich danke daher den Leipziger Projektleitern Susanne Wagner und Christian Glück, dass sie es mir ermöglicht haben, mich in den letzten Wochen vorwiegend auf die Abgabe dieser Arbeit zu konzentrieren.

„Kein Weg ist lang, mit einem Freund an der Seite." Ich danke meinen Freunden von ganzem Herzen, dass sie mich auf diesem Weg geduldig begleitet, motiviert und bestärkt haben. Mein größter Dank gilt meinem Mann Thilo, für seine liebevolle Unterstützung und das unendliche Verständnis. Danke, Noemi und Thilo, für das Lachen, das Weinen und die Liebe.

I SYNOPSIS

1 Introduction

Prosody is a rich source of information in spoken language comprehension. Encompassing the suprasegmental properties of speech—such as intonation, rhythm, timing, and intensity—the prosodic structure conveys various aspects of an utterance or a speaker's intention: It marks types of utterance such as questions and statements and is used to emphasize contrasting or new information and to focus on relevant utterance contents. Moreover, it indicates sarcasm or the use of irony, signals turn-taking in conversation, and may reveal the speaker's emotional state or attitude. In tonal languages (such as Mandarin Chinese, which is the most widely spoken tonal language), prosody additionally has a strong lexical function in that pitch information is used to encode lexical or grammatical meaning differences.

Above and beyond the aforementioned linguistic and communicative functions, the pivot role of prosodic information is to divide the continuous speech stream into chunks, so that each chunk sounds like a complete, self-contained section of the utterance. This *chunking* or *prosodic phrasing* plays a critical role in language processing (e.g., Cutler, Dahan, & van Donselaar, 1997; Frazier, Carlson, & Clifton, 2006) and spoken language acquisition (e.g., Gervain & Mehler, 2010; Morgan & Demuth, 1996; Speer & Ito, 2009), because words are never arbitrarily chunked into prosodic phrases, but instead the prosodic structure reflects the semantic and syntactic properties of an utterance. Extra-linguistic factors (such as speech planning or breathing) rarely interfere with this grouping of coherent utterance parts. Thus, prosodic phrasing helps to convey the intended meaning of a string of words.

Specifically, infants as well as adults benefit from a close syntax-prosody mapping, that is, the edges of major prosodic phrases usually coincide with syntactic boundaries (e.g., Downing, 1970; Selkirk, 2005; for German: Truckenbrodt, 2005). Therefore, prosodic phrasing information enables the listener to draw inferences about the underlying syntactic structure and hence largely contributes to auditory sentence comprehension. Accordingly, numerous behavioral studies demonstrated an influence of prosodic phrasing on syntactic analysis (for comprehensive reviews, see Cutler et al., 1997; Wagner & Watson, 2010).

In line with its central role in auditory language comprehension, prosodic phrasing is also an important source of information in language acquisition, because it enables infants to segment the continuous speech signal and to extract linguistically relevant units. According to the concept of *Prosodic Bootstrapping* (Gleitman & Wanner, 1982; Morgan & Demuth, 1996; see also Höhle, 2009), prosodic information like stress, rhythm, and intonation helps young learners to acquire syntactic and lexical knowledge of their native language. Correspondingly, young infants have been shown to be highly sensitive to prosodic boundary information and to use it for segmentation of clauses or words (for a review, see Speer & Ito, 2009).

While the importance of prosodic phrasing for language acquisition and auditory language processing is evident, it is not yet well understood how the acoustic correlates of prosodic boundaries are processed in order that infants and listeners can interpret them linguistically. In other words, when a listener actually perceives a prosodic boundary as such, the language processing system must have interpreted acoustic features occurring in the speech stream as signaling the closure of a prosodic phrase. Hence, successful prosodic phrasing depends on the on-line classification of acoustic properties such as changes in the fundamental frequency, segmental duration, and pause occurrence. To date, little is known about listeners' cognitive processes for making sense of these cues in linguistic interpretation. The current work thus seeks to shed light on the procedural mechanisms underlying prosodic boundary perception. In particular, it tackles the following questions:

- How are prosodic boundary cues processed? Under which conditions is a prosodic boundary perceived as such?
- What is the relation between the acoustic correlates encountered and the perception of a prosodic boundary? Considering the set of acoustic features known to correlate with prosodic boundaries, which of them are relevant to cueing boundary perception?
- When and how does the ability to make use of prosodic boundary cues develop during infancy?

To answer these questions, this dissertation deals with two factors considerably affecting boundary perception: (1) contextual influence in terms of the phrase length preceding the boundary and (2) the influence of specific boundary cues or cue combinations on infant and adult boundary perception. It encompasses three independent studies. Study I (Chapter 7) explores whether adults' perception of a prosodic boundary is affected by its position within an utterance. It sheds light on contextual effects on boundary perception, commonly referred to as an effect of phrase length, by systematically varying the amount of prosodic context while another factor, that is, the acoustic correlates signaling a prosodic boundary, was kept constant. In contrast, in Study II (Chapter 8) and Study III (Chapter 9) the boundary position was kept constant, but the concomitance of boundary cues was systematically varied by means of acoustic manipulation. While a considerable amount of research was aimed at describing the acoustic correlates of a prosodic boundary and thus identifying the prosodic cues assumed to trigger boundary perception (see 2.2), little is known about the impact of specific cues or cue combinations and their interplay in perception. Hence, the current work puts special emphasis on the systematic acoustic manipulation of prosodic cues to investigate their specific role in boundary perception. The studies were carefully designed to infer the role of specific boundary cues and cue combinations both in adult boundary perception (Study II) and in the development of boundary perception by testing of infants from two age groups (Study III).

Following this brief outline, the remaining part of the synopsis is divided into five sections, each of them addressing core issues that underlie or result from the empirical work presented in the peer-reviewed journal articles: The theoretical considerations (Chapter 2) cover both the theoretical background regarding the concept of prosodic phrasing as well as the relevant previous research on boundary perception. In the methodological considerations (Chapter 3), I discuss two outstanding characteristics regarding the experimental set-up, namely the application of the event-related potential (ERP) technique to study boundary perception and the nature of the stimulus material used throughout the five ERP experiments included in Study I to III. The research questions are spelled out in the subsequent Chapter 4 that also refers to the main arguments leading to hypotheses and possible implications of the results. Finally, a summary of the main findings is provided (Chapter 5); in the conclusions (Chapter 6), their significance is discussed in reference to the research questions and hypotheses.

2 Theoretical considerations

The current work is based on theoretic assumptions regarding prosodic phrasing and on findings from previous research characterizing the acoustic correlates signaling a prosodic boundary. This fundament is outlined in the following sections. Subsequently, previous studies on the perception and weighting of prosodic boundary cues both in adults and infants are reviewed, while the closing section covers the question of contextual influence on prosodic boundary perception.

2.1 Prosodic phrasing

In the classic theory of suprasegmental phonology, it is assumed that utterances are chunked into prosodic constituents of different strength that are hierarchically organized (see, e.g., Nespor & Vogel, 1986; Selkirk, 1986) as reflected in(1):

(1) The prosodic hierarchy (following Selkirk, 1986):

Utterance (U)
|
Intonational Phrase (IP)
|
Phonological Phrase (PhP)
|
Prosodic Word (PW)
|
Foot (Ft)
|
Syllable (σ)

According to the strict layer hypothesis (Selkirk, 1984, 1986), a prosodic unit of one level of the prosodic hierarchy is exhaustively parsed into constituents of the next-lower level. Moreover, the assumption of strict layering interdicts skipping of one layer of representation (e.g., an Intonational Phrase cannot be parsed into a Phonological Phrase and a Prosodic Word) and recursivity (e.g., an Intonational Phrase cannot consist of a Phonological Phrase and another Intonational Phrase).

Intonational and Phonological Phrases are pivotal constituents with regard to the syntax-phonology interface, as their boundaries typically fall together with phrasal and clausal syntactic boundaries, respectively (e.g., Downing, 1970; Selkirk, 2005;

for German: Truckenbrodt, 2005). Due to this mapping between prosodic and syntactic units, prosodic boundary perception can be an important guide to the syntactic structure of spoken language and thus contributes to sentence comprehension and helps in acquisition.

However, the relationship between prosodic and syntactic structure is inconsistent, as there is no strict association between a specific prosodic constituent and a syntactic phrase (see, e.g., Cole, 2014). For instance, the postulated set of categorical prosodic representations falls short of complying with production data where different boundary strengths cannot be interpreted as categorical phonological differences between boundary types (e.g., recursive phrasing found in coordinate structures; Wagner, 2005). A related problem concerning the assumption of distinct prosodic categories is that a phonetic definition of different types of prosodic phrase categories is inconclusive, as different types of phrases (i.e., Phonological and Intonational Phrase) often look phonetically similar. In other words, when boundaries of different strengths can be discerned, the differences are quantitative rather than qualitative.

Attempts to overcome this mismatch between syntactic and prosodic structure involved postulating additional ad hoc categories (e.g., minor vs. major Phonological Phrases, Selkirk & Tateishi, 1988; Intermediate vs. Accentual Phrase, Beckman & Pierrehumbert, 1986) and, more recently, the admission of recursive prosodic domains (Selkirk, 2005; Truckenbrodt & Féry, 2015), or—as an alternative to the prosodic hierarchy—a recursive prosodic system (Selkirk, 2011; Wagner, 2005, 2010). In particular, beyond the categorical distinctions, Intonational Phrases of different strengths should be distinguishable (see, for a discussion, Ladd, 2008).

The ongoing theoretical debate regarding the nature of prosodic representations and their relation to syntactic structure (see, for a discussion, Wagner & Watson, 2010) notwithstanding, the prosodic structures examined in this work are classified as Intonational Phrases and the respective boundaries are referred to as Intonational Phrase boundaries. This classification is not uncontroversial, given the syntactic nature of the experimental stimuli, namely coordinate structures instead of clauses (see 3.2). However, bearing in mind that Intonational Phrases are not exclusively aligned to clausal units (e.g., Watson & Gibson, 2004), the classification arises from the prosodic disambiguation instantiated in the current material (see below). Moreover, the interested reader is referred to Petrone et al. (in press) for a detailed evaluation of the prosodic structures under investigation.

I will use the term *major prosodic boundary* (PB; also spelled out as “prosodic break”, see Bögels, Schriefers, Vonk, & Chwilla, 2011) in what follows, given that a fine-grained differentiation between higher level prosodic constituents such as Phonological or Intonational Phrase is clearly beyond the scope of the present work. Rather than representing a distinct prosodic category, a PB is defined by its acoustic correlates occurring uttered speech signal. These acoustic correlates serve as cues to (major) prosodic boundaries and are described in the following section. Further on, the terms *Intonational Phrase boundary* (IPB) and *major prosodic boundary* (PB) will be used interchangeably to denote the prosodic phrase boundaries under investigation.

In sum, the hierarchical organization of prosodic phrases shows a systematic relation to the lexico-semantic and syntactic structure, in that prosodic constituents of different strengths are aligned with lexical and syntactic units. Major prosodic boundaries thus play a crucial role in both sentence comprehension language acquisition, allowing to infer the edges of syntactic phrases and clauses.

2.2 Prosodic boundary cues

The edges of prosodic phrases are phonetically marked. Across languages, various acoustic phenomena pertaining to the domains of timing, fundamental frequency, voice quality, and intensity are known to signal a PB (see, for a review, Cole, 2014; Wagner & Watson, 2010). Despite a considerable range and interspeaker as well as cross-linguistic variability in the use of acoustic correlates, the current work addresses the three main boundary cues identified to signal a PB, namely *pause, final lengthening,* and *pitch change* (for German see, e.g., Gollrad, Sommerfeld, & Kügler, 2010; Kentner & Féry, 2013; Kohler, 1983; Peters, Kohler, & Wesener, 2005; Petrone et al., in press).

A *pause*, that is, the insertion of an interval of silence at prosodic phrase junctures, and an increased duration of the segments immediately preceding the boundary (i.e., *final lengthening*) are closely related durational PB cues. They have been argued to contribute to a single percept of a break (see Wagner & Watson, 2010) as listeners may perceive a pause even if no interval of silence, but only lengthened phrase-final segments are presented (e.g., Martin, 1970). Still, taking into account that the acoustic correlates are ascribed diverging distributions and linguistic functions, several experimental studies treat pause and final lengthening as two boundary cues, contributing independently to PB perception (see below). Regarding the dimension of fundamental frequency and its perceptual correlate pitch, *pitch change* (i.e., a pitch rise or fall) clearly signals a PB by indicating the presence of a boundary tone. Pitch change as a reflection of a boundary tone is usually followed by a pitch reset at the left edge of the following phrase.

While the use of the prosodic cues is a universal phenomenon, the cue weighting, that is, the relative importance of one cue (e.g., final lengthening) as compared to another one (e.g., pitch change), assumedly develops as a function of exposure to the ambient language and may thus differ across languages (see 2.2.2 below).

Regarding the distribution of prosodic boundary cues in German spontaneous speech, a corpus study (Peters et al., 2005) found that although cue combinations are frequent in production, 31.5% of all turn-internal boundaries were marked by only one of the three phrase boundary cues (20.8% pitch change, 9.4% final lengthening, 1.3% pause). Moreover, the combination of pitch change and final lengthening occurred predominantly among the possible combinations of two boundary cues: While 24.6% of all boundaries were marked by the combination of pitch change and final lengthening, only 8.4% were marked by the co-occurrence of final lengthening and pause, and only 4.9% by pitch change and pause.

Under the assumption that prosodic cue weighting in perception depends on—or is, at least, mutually influenced by—the distribution of the cues in the ambient language, the corpus data indicates that German PB perception does not necessarily require the presence of all three boundary cues. Further, it suggests that pitch change, final lengthening, and pause differ with regard to their role in PB perception in that specific cues or cue combinations have a larger impact than others. However, research regarding the contribution and the relative importance of prosodic boundary cues is sparse, especially concerning German PB perception, as will be outlined in the following section.

2.2.1 Adult processing of prosodic boundary cues

The contribution of specific acoustic correlates to the perception of a PB has predominantly been studied in English language comprehension. In a pioneering study, Streeter (1978) used ambiguous algebraic expressions such as "(A plus E) times O" and "A plus (E times O)" that were acoustically manipulated by interchanging the pitch contour and the duration pattern (i.e., final lengthening) of the two alternative phrasings. Participants were asked to render a grouping decision by stating which of the above cited bracketings was conveyed by the speaker. Both, duration pattern and pitch contour, provided as single cues affected the listeners' judgments in that there was a higher tendency to opt for the alternative phrasing when one of the cues was manipulated. However, only the combination of both cues completely shifted the listeners' decision from the original phrasing to the alternative phrasing. Therefore, Streeter (1978) assumes an additive effect of the two prosodic cues.

Unlike Streeter, who studied the role of pitch change and final lengthening, Scott (1982) and Aasland and Baum (2003) focused on the impact of the two temporal cues—final lengthening and pause—in a phrasal grouping task. Scott (1982) used short English sentences containing an ambiguous string of three names separated by different conjunctions (e.g., "Kate and Pat or Tony will come"), while Aasland and Baum (2003) employed the phrase "pink and black and green". The natural stimuli were edited in a stepwise manner by systematically increasing or decreasing the amount of final lengthening and/or pause duration at the boundary position.

Scott (1982) found the combination of final lengthening and pause duration to be an effective boundary cue, increasingly shifting the listeners' decision towards the alternative phrasing. However, an extended pause provided as single cue was found to be as effective as the combined occurrence of final lengthening and pause. The impact of final lengthening as a single cue was not tested. Aasland and Baum (2003) confirmed that the combined occurrence of final lengthening and pause effectively triggers PB perception. Crucially, also a sufficiently large final lengthening cue shifted the participants' decision towards the alternative phrasing, whereas—in contrast to the findings by Scott (1982)—even the longest pause cue was insufficient to provoke PB perception in the absence of final lengthening.

Presumably, the contradictory results can be traced back to the size of the presented cues: The pause duration continuum for the manipulated stimuli in Scott (1982) ranged from 0 ms up to 562 ms, whereas the longest pause cue used by Aasland and Baum (2003) was only 160 ms long: Although pauses up to 256 ms had been found in the natural speech stimuli, the experimental pause duration was manipulated only along a 5-step-continnum with a step size of 40 ms. Thus, the diverging results concerning the impact of pause as a single cue may reflect confounding differences in the compilation of the stimulus material.

The same reasoning holds true for inconsistent results recently obtained for cue weighting in Mandarin Chinese: Zhang (2012) compared the perception of pitch change, final lengthening, and pause cues in English and in Mandarin Chinese. In both languages, listeners' grouping decisions reflected that final lengthening had the smallest impact on PB perception. The remaining two cues were weighted differently: English listeners relied more on pause than on pitch contour change, while Chinese listeners' decisions were more influenced by pitch manipulation than by pause. The stronger reliance on pitch information was associated with the phonemic status of pitch in the Chinese phonological system as compared to non-tonal languages like English. In contrast, Yang, Shen, Li, and Yang (2014) claimed pause to be the most powerful PB cue in Mandarin Chinese, while pitch and final lengthening were assumed to be weaker and, in addition, perceptually equivalent cues. The pause duration employed in this study represented the duration observed in natural speech and averaged out at 270 ms, whereas Zhang (2012) inserted pauses that lasted at maximum 80 ms, although pauses of more than 300 ms occurred in naturally produced stimuli with PB.

Two recent studies addressed the perception of prosodic boundary cues in German. Gollrad (Gollrad, 2013; Gollrad et al., 2010) investigated the role of prosodic boundary cues in the resolution of syntactic case ambiguities. Two different syntactic structures were examined in order to distinguish between Phonological Phrase (here: occurring clause-internal) and Intonational Phrase (here: sentence-final) boundaries (see 2.1). Initially, Gollrad found for clause-internal boundaries that durational cues (i.e., combining final lengthening and pause) are perceptually more relevant than cues pertaining to the pitch contour (Gollrad et al., 2010). From subsequent experiments aimed at disentangling the contribution of pause and final lengthening, the conclusion was drawn that final lengthening had the largest impact on clause-internal case ambiguity resolution. In contrast, as for sentence-final boundaries, Gollrad (2013) found pitch information to be most decisive, indicating that boundary tone information plays a major role in the perception of a PB.

Closely related to the current work, Petrone et al. (in press) studied prosodic phrasing in German coordinate structures. As for the perception of prosodic boundaries, three behavioral phrasal grouping experiments were carried out, each focusing on one of the three main boundary cues. To that end, stimuli were used in which the respective cue was acoustically manipulated in a stepwise manner, while the other two cues were neutralized. Results were interpreted in terms of individual contributions to PB perception, since for each single cue the acoustic manipulation affected listeners' prosodic judgments and shifted their responses towards the alternative phrasing. Notably, this shift in the grouping decisions was rather abrupt for the pause cue, but occurred gradually in response to pitch change and final lengthening cue manipulations. Hence, listeners exploit the pause cue in a categorical fashion, while pitch change and final lengthening constitute rather gradual cues to prosodic boundaries.

Apparently, methodological differences (including cue size, nature of the acoustic manipulations, and experimental task) limit the conclusions that can be drawn from the cited work that partly yields inconsistent results as to the relative importance of the prosodic cues under investigation. However, several phrasal grouping studies suggest that pause is a very salient boundary and seems to be a decisive PB cue, at least in non-tonal languages. The pre-dominance of the pause cue is further supported by Peters (2005) who found that the presence of a pause can mask effects of other boundary cues in German boundary perception. Moreover, there is consistent evidence that both final lengthening and pitch change contribute to the perception of a PB. Yet, their specific role and effectiveness as single cues or in absence of the optional pause cue remains unclear.

For this reason, the present work focuses on the contribution of pitch change and final lengthening to German phrase boundary perception, both as single cues and in combination. Abstracting from the apparently salient pause cue bears the advantage that light is shed on the two other, acoustically more subtle cues. In particular, it allows to tease apart the role of final lengthening from the impact of the pause. In addition, this thesis takes advantage of on-line measures (ERPs, see Chapter 3.1), tracking the immediate impact of the prosodic cues during incremental boundary processing. Since ERPs allow language perception studies independent of specific task demands or an overt response performance, the use is all the more beneficial as Study III investigates boundary cue processing in infants. The respective background—sketching infants' abilities to make use of prosodic information—is given below.

2.2.2 Infant processing of prosodic boundary cues

One of the core aspects of the prosodic bootstrapping account (Gleitman & Wanner, 1982) is that an early sensitivity to prosodic boundaries facilitates the detection of syntactic units such as clauses and syntactic phrases. Language acquisition research concordantly demonstrates infants' sensitivity to prosodic information signaling clausal and phrasal boundaries (e.g., Hirsh-Pasek et al., 1987; Jusczyk et al., 1992; see, for a review, Speer & Ito, 2009).

Moreover, several behavioral studies focused on infants' use of prosodic boundary information for speech segmentation and clause recognition. In particular, Nazzi, Kemler Nelson, Jusczyk, and Jusczyk (2000), Soderstrom, Kemler Nelson, and Jusczyk (2005), and Seidl (2007) tested six-month-old English-learning infants and investigated their ability to recognize previously heard word sequences embedded in continuous speech using the head-turn preference procedure (HPP, Hirsh-Pasek et al., 1987; Jusczyk & Aslin, 1995). The studies employed similar crossed designs in which each of two groups of infants was familiarized with a word sequence presented in two prosodic versions: One was prosodically coherent, that is, the critical sequence formed a clause and occurred as a single prosodic phrase, while the other one was incoherent in that the sequence was a non-clausal unit that contained parts of two prosodic phrases and spanned a prosodic boundary (see Table 2.1 for exemplar stimuli). During the test phase infants heard text passages that either contained the prosodically coherent or the incoherent word sequence. Importantly, the test passage containing the coherent word sequence heard by the first group also contained the incoherent sequence used during familiarization of the second group of infants and vice versa (see Table 2.1).

Table 2.1 | Exemplar stimuli presented during familiarization and test phase (adapted from Seidl, 2007).

Familiarization		
Familiarization groups	Clause	Non-clause
Group A	[leafy vegetables taste so good]	[leafy vegetables][taste so good]
Group B	[rabbits eat leafy vegetables]	[rabbits eat][leafy vegetables]
Test phase (passages presented to both familiarization groups)		
I	John doesn't know what **rabbits eat.** ***Leafy vegetables*** *taste so good*. They don't cost much either.	
II	Many animals prefer green things. **Rabbits eat** ***leafy vegetables.*** *Taste so good* is rarely encountered.	

The studies concordantly yielded a preference pattern that depended on the familiarization: Infants from both groups preferred to listen to the text passage that contained the prosodically coherent word sequence they heard during familiarization. This shows that infants recognize a word sequence better when it is a prosodically coherent clausal unit and that they recognize the sequences on the basis of the available prosodic boundary cues. Hence, the findings of these studies do not only show that six-month-old infants are sensitive to prosodic phrasing information, they further suggest that infants use this information to segment the speech stream and to recognize clausal units.

Recent language acquisition research examined the development of a language-specific weighting of prosodic cues. A core question is whether infants need all three prosodic cues to perceive a PB or whether and at which age a subset is sufficient to perceive a PB in the ambient language. To investigate the prosodic cue weighting in English-learning infants, Seidl (2007) altered the difference between sequences presented in the familiarization phase (see Table 2.1) by neutralizing either one or two of the prosodic cues. She found that neither the absence of final lengthening nor of the pause cue kept six-month-old English learners from recognizing the phrases in continuous speech. In contrast, when the pitch cue was neutralized, the infants no longer demonstrated successful clause recognition. This pattern of results suggests that by six months, English-learning infants are able to perceive a PB on the basis of a subset of the three main boundary cues, as long as the pitch cue is present.

An extension of the study testing younger English-learning infants (Seidl & Cristià, 2008) revealed that four-month-olds—in contrast to the six-month-olds—need all three prosodic cues for PB perception. This difference between the age groups is interpreted in terms of a development in cue reliance, reflecting an attunement of speech perception to the ambient language and presumably turning boundary cue processing into an adult-like perception pattern. In order to investigate infants' boundary cue weighting in a language other than English, the experimental design was adapted to Dutch (Johnson & Seidl, 2008). Crucially, the pause cue proved to be necessary for PB perception in six-month-old Dutch learners. As this finding contrasts the results found for English learners, the diverging results are supposed to display cross-linguistic variation, suggesting that at six months of age, infants' prosodic cue weighting may already be affected by the ambient language.

Further evidence for an early development in PB perception stems from two recent HPP studies with German six- and eight-month-old infants (Wellmann, Holzgrefe, Truckenbrodt, Wartenburger, & Höhle, 2012, Wellmann, Holzgrefe-Lang, Truckenbrodt, Wartenburger, & Höhle, submitted). Here, the stimulus material consisted of short sequences of three coordinated names (similar to the stimuli used in the present work, see Chapter 3.2). The sequences either formed a single major prosodic phrase without internal PB, or they were made up by two phrases with an internal PB. Eight-month-olds successfully discriminated the two prosodic patterns not only when natural stimuli with fully marked PBs were employed, but also for acoustically manipulated stimuli containing PBs that were only signaled by pitch change and final lengthening. However, the stimuli could not be discriminated when the PB was solely signaled by either pitch change or by final lengthening (Wellmann et al., 2012). In contrast, for German six-month-olds, the combination of pitch and lengthening was not sufficient (Wellmann et al., submitted); they needed a pause cue to discriminate the prosodic patterns. Similarly to the results for English, this indicates a developmental change in German infants' PB perception and is interpreted in terms of a shift from rather basic acoustic detection at six months that heavily relies on the pause cue towards a more sophisticated linguistic processing of PB cues at eight months.

In sum, behavioral studies have demonstrated that infants are highly sensitive to prosodic boundaries and use this prosodic information for speech segmentation, at least within the second half of the first year of life. Moreover, results point to a development regarding the infantile ability to make use of specific prosodic cues and cue combinations for PB perception. Taking into account studies from three different languages (English, Dutch, and German) suggests that this developmental path reflects an attunement to the ambient language. However, neurophysiological research on infantile boundary processing is sparse and has yielded inconclusive results (see Chapter 3). Therefore, Study III investigates on-line processing of prosodic boundary cues in infants using stimuli that allow a direct comparison to adult cue perception (Study II) and to behavioral results from German learning infants (Wellmann et al., 2012; Wellmann et al., submitted). A basic prerequisite to investigate boundary perception using acoustically manipulated stimuli is to evaluate the appropriateness of the original natural stimulus material. Hence, the preceding Study I employed comparable natural stimuli containing a PB at varying utterance position. It is aimed at exploring the influence of prosodic context on PB perception, a crucial factor that is outlined in the following section.

2.3 Contextual influence on boundary perception

Numerous behavioral studies on adult spoken language comprehension demonstrated an influence of prosodic phrasing on syntactic analysis (see, for a review, Cutler et al., 1997 and, among others, Kjelgaard & Speer, 1999; Price, Ostendorf, Shattuck-Hufnagel, & Fong, 1991; Sanderman & Collier, 1997; Schafer, 1997). However, nature and time-course of this interaction are not yet well-understood. Although the prosody-syntax-mapping relies on the close alignment of the edges of prosodic and syntactic units, prosodic information is not only present locally at the boundary position. Instead, it unfolds globally throughout an utterance. For instance, Clifton, Carlson, and Frazier (2002) argue against a solely local interpretation of prosodic boundaries. They provide evidence that the prosodic context in which a PB occurs plays a major role, as the listener determines on a contextual basis whether a prosodic boundary is relevant for syntactic parsing decisions or not: The occurrence and the strength of a preceding boundary as well as the length of the preceding phrase affected PB perception, irrespective of the absolute strength of the PB. More recent evidence for this impact of relative PB strength on boundary perception is provided by Snedeker and Casserly (2010) as well as Wagner and Crivellaro (2010).

Phrase length, that is the amount of material processed within one constituent, has been shown to affect prosodic phrasing in speech production and in the processing of implicit prosody in silent reading (Fodor, 1998; Gee & Grosjean, 1983; Watson & Gibson, 2004; Hwang & Schafer, 2009). Regarding the perception of prosodic boundaries, Clifton, Carlson, and Frazier (2006) demonstrate that phrase length affects the comprehension of syntactically ambiguous sentence structures. These findings strongly suggest that globally distributed prosodic information is integrated into the processing of prosodic boundaries as markers of syntactic structure. However, based on the behavioral data it cannot be decided whether the global prosodic structure has a direct impact on the perception of prosodic boundaries or whether the observed effects occur later during the process of sentence interpretation. Thus, on-line methods like the ERP technique (see below) are needed to identify whether incremental prosodic processing leads to an immediate impact of the prosodic context on PB perception. The application of ERPs to study effects of prosodic context and boundary cue manipulation on PB processing will be further illustrated in the next section.

3 Methodological considerations

3.1 Using ERPs in the study of prosodic boundary perception

Behavioral studies make use of so-called off-line methods (e.g., judgment and reaction time data, or unwitting reactions like head-turns or sucking rates in infants) that yield the result of a complex perceptual operation. In addition, they require participants to perform a task (or, in case of unwitting reactions, at least a motor response) which involves several processing steps until the result can be measured. This result is usually obtained only after the whole stimulus has been processed. Even when behavioral measures are tried to be obtained instantly (e.g., via keystroke during stimulus presentation), the responses have a low temporal resolution, that is, they do not offer insights as to how and when processing has been affected by the experimental manipulation. This is because off-line measures are not obtained continuously, but only at one specific point in time (i.e., within or after stimulus presentation).

The current work, however, sets out to study on-line perception of prosodic boundaries by means of a well-established neurophysiological research method, namely the event-related potential (ERP) technique (see, e.g., Luck, 2014 for a comprehensive overview). In short, ERPs are „electrical *potentials* generated by the brain that are *related* to specific internal or external *events*" (Luck, 2012, p. 523, emphasis in original). They are obtained non-invasively by continuously recording participants' electroencephalogram (EEG) while the experimental stimuli are presented. Subsequently, a mean is taken for all trials of one experimental condition to extract the ERPs specifically related to the presented stimuli by averaging out non-related, random brain activity. An observed ERP waveform contains several so-called ERP components that can be described in terms of the brainwave's morphology (i.e., polarity, amplitude, and latency) and with regard to the scalp topography (i.e., the spatial distribution across electrode sites). Further, ERP components are replicable and sensitive to a specific experimental manipulation—and thus reflect specific underlying neural processes (see Luck & Kappenman, 2012 for a comprehensive discussion and a detailed review of the major ERP components). Therefore, ERP studies usually focus on specific ERP components to address questions regarding the neurocognitive processing architecture.

Notably, using ERPs allows language perception studies independent of specific task demands or an overt response performance (see Männel & Friederici, 2008). Moreover, since the ERP technique discloses the electrical brain activity directly resulting from a specific event or stimulus, it captures the immediate impact of (linguistic) information during on-line processing. Due to this high temporal resolution, it is a useful tool to address questions on the integration of prosodic information. Accordingly, studies on sentence processing using ERPs demonstrate an early integration of prosodic information and thus an immediate influence on the syntactic parsing process (see, e.g., Eckstein & Friederici, 2006). In particular, ERPs can be employed to study the time course and underlying mechanisms of PB processing (see, for a review, Bögels et al., 2011), since a *Closure Positive Shift* (CPS, see below) signals the processing of a major PB perceived in the stimulus material.

Throughout this thesis, the CPS and the advantages of the ERP technique are used to contribute to our understanding of prosodic boundary processing: First, in Study I the high temporal resolution allows to investigate the immediate impact of contextual prosodic information (i.e., phrase length) on the perception of prosodic boundaries. Then, Study II benefits from the combined use of off-line and on-line research techniques, namely a prosodic judgment task and ERP recordings, in that the complementary results allow conclusions to be drawn on a methodological refinement. Finally, Study III on infants' processing of PB cue makes use of the ERP technique by directly and task-independently investigating a population that, using behavioral measures, could be assessed only very indirectly (e.g., via head-turn orientation times).

3.1.1 The Closure Positive Shift (CPS)

Steinhauer, Alter, and Friederici (1999) conducted an ERP study comparing the perception of German sentences that contained either one or two PBs. Brain responses to prosodic violations (here, a prosodic boundary inserted at a non-boundary position) indicated that syntactic processing was misled by prosodic information at an early processing stage. Crucially, in response to each PB a broad positive deflection occurred in the ERPs, which was broadly distributed over the scalp, but found to be largest at central and parietal electrode sites. Due to the co-occurrence with the closure of a PB, the component was termed *Closure Positive Shift* (CPS). To date, the CPS is a well-established indicator of adult PB perception in various languages: in German (e.g., Männel & Friederici, 2009; Pannekamp, Toepel, Alter, Hahne, & Friederici, 2005), English (e.g., Itzhak, Pauker, Drury, Baum, & Steinhauer, 2010; Pauker, Itzhak, Baum, & Steinhauer, 2011; Peter, McArthur, & Crain, 2014), Dutch (e.g., Bögels, Schriefers, Vonk, Chwilla, & Kerkhofs, 2010; Kerkhofs, Vonk, Schriefers, & Chwilla, 2007), Swedish (Roll & Horne, 2011), Japanese (Wolff, Schlesewsky, Hirotani, & Bornkessel-Schlesewsky, 2008), Mandarin Chinese (Li & Yang, 2009), and Korean (Hwang & Steinhauer, 2011).

Regarding the nature of the CPS, Pannekamp et al. (2005) provided evidence for a truly prosodic origin by demonstrating that the CPS can be elicited in the absence of semantic, syntactic, or segmental information, albeit with varying scalp distributions. Moreover, the CPS has been elicited in the absence of the pause cue (e.g., Männel & Friederici, 2009; Steinhauer et al., 1999). This clearly shows that the CPS is not a variant of an obligatory onset component signaling the detection of new auditory input (e.g., after a period of silence). Thus, although the CPS may partly overlap with so-called *early auditory evoked potentials*, it substantially differs from this indication of lower level processing related to speech material following the PB. Moreover, with regard to the impact of specific prosodic boundary cues, this suggests that a pause is rather optional and not mandatory for PB perception.

Another argument against the CPS reflecting low level acoustic processing is that the CPS can also be elicited during silent reading (e.g., at comma positions; Drury, Baum, Valeriote, & Steinhauer, 2016; Steinhauer, 2003; Steinhauer & Friederici, 2001; but see also Kerkhofs, Vonk, Schriefers, & Chwilla, 2008), indicating that implicit prosodic phrasing plays a role in written language processing. In a silent reading study on Korean, Hwang and Steinhauer (2011) found that only longer sentence-initial constituents elicited a CPS, while no effect was found for short subject noun phrases. Here, the CPS was considered to reflect the subvocal generation of an additional prosodic boundary, which was only triggered by long constituents. Notably, the study demonstrates an influence of constituent length on implicit prosodic phrasing and thus indicates that the CPS may be used to study the direct impact of contextual (prosodic) information on prosodic phrasing.

In sum, there is compelling evidence that the CPS component reflects prosodic boundary perception in auditory language processing, as the occurrence of the CPS depends on prosodic information in the speech signal, that is, the prosodic cues marking the closure of a PB, but not on other linguistic information or on mere acoustic changes in the input. Furthermore, the occurrence of the CPS during silent reading is supposed to reflect covert prosodic processing and indicates a sensitivity to contextual effects on implicit prosodic phrasing.

3.1.2 The Infant CPS

To date, there is mixed evidence whether the CPS can also be found in infants or whether it emerges only later in development. Pannekamp, Weber, and Friederici (2006) tested eight-month-old German-learning infants and found a positive shift in response to sentence stimuli containing an internal PB (versus stimuli without internal PB). Although this shift was delayed as compared to the adults' responses (cf. Pannekamp et al., 2005, Steinhauer et al., 1999), the authors considered it as an infant CPS reflecting their ability to perceive prosodic boundaries.

Männel and Friederici (2009) intended to replicate the infant CPS for five-month-old German-learning infants. Results showed indeed a positive shift in response to the condition containing an internal PB. However, detailed data analyses revealed that the effect could not be clearly attributed to the processing of the PB. Instead, it seemed to derive at least partially from a so-called obligatory onset component (see, e.g., Kushnerenko et al., 2002). This component signals low-level processing of newly incoming acoustic input following an interval of silence, that is, after the pause at the prosodic boundary under investigation. This would also explain the delayed onset of the positive shift found by Pannekamp et al. (2006): Since the obligatory onset component is a response to the acoustic input after the pause, it usually starts later than the CPS signaling the perception of prosodic boundary cues (though there may be a considerable overlap, see adult studies, above).

Testing of older children (Männel & Friederici, 2011) indicated that only three- and six-year-olds—but not 21-month-olds—showed an adult-like CPS in response to fully marked PBs. Hence, Männel and Friederici (2009, 2011) questioned the existence of an infant CPS and allocated the emergence of the CPS to the acquisition of sophisticated syntactic knowledge gained within the third year of life. Regarding the role of specific boundary cues, Männel, Schipke, and Friederici (2013) found that six-year-olds show an adult-like CPS when exposed to PBs marked by pitch change and final lengthening only, while 3-year olds needed the combination of final lengthening and a pause (Männel & Friederici, 2016). This finding was interpreted in terms of a larger impact of durational cues (i.e., the combination of final lengthening and a pause) on German toddlers' PB processing than pitch cues, which is in line with results obtained for German adults (Gollrad et al., 2010).

Still, it remains unclear how this neurophysiological development in toddlers and preschoolers relates to the remarkable competencies that infants possess in terms of an early sensitivity to prosodic phrasing information (e.g., Hirsh-Pasek et al., 1987; Jusczyk et al., 1992). Behavioral studies clearly show that already six-month-old infants use PB cues to segment the speech input and to extract meaningful units (e.g., Gout, Christophe, & Morgan, 2004; Nazzi et al., 2000; Seidl, 2007; Soderstrom, Seidl, Kemler Nelson, & Jusczyk, 2003; Shukla, White, & Aslin, 2011). It seems reasonable to assume that this early use of prosodic boundary cues for segmentation is also reflected in infants' brain responses. Moreover, this raises the question whether the developmental shift in PB cue weighting observed within the first year of life (Seidl, 2007; Seidl & Cristià, 2008; Wellmann et al., 2012; Wellmann et al., submitted) can also be evidenced on the neurophysiological level.

3.2 Using coordinate structures

In contrast to previous ERP studies using long sentence material, the current work takes advantage of stimulus material that consisted of lists of coordinated names. More precisely, in each experiment the participants were presented lists of three German names conjoined with *und* ('and') or *oder* ('or'), as illustrated in the following examples (bracketing indicates the alternative internal groupings, the respective PB position is marked by a hash mark):

(2) a. (Mona) # (oder Lena und Lola)
b. (Mona oder Lena) # (und Lola)

(3) a. (Moni und Lilli und Manu)
b. (Moni und Lilli) # (und Manu)

The stimuli were syntactically ambiguous in that a possible subgrouping of the three names depended on the presence and position of an internal PB: In Study I, the stimuli contained a PB either after the first or after the second name as in (2), whereas in Study II and III, the stimuli either contained no internal PB or a PB after the second name (3) that was signaled by different prosodic cues or cue combinations, depending on the experimental condition.

Although the current stimulus material deviates from the sentence material commonly used in ERP studies on PB perception (see 3.1.1), using coordinated lists bears several advantages. First, the limited range of linguistic material facilitates a careful control of the phonetic and phonological characteristics of the material. For instance, only disyllabic, trochaic names were used. The names were composed of four sonorants to allow a thorough acoustic analysis and—in case of Study II and III—manipulation of the stimulus material. Second, the simple structures facilitate a fine-grained temporal analysis of the material which is of special importance in ERP studies: Due to the high temporal resolution, stimuli from different conditions usually need to be temporally aligned (i.e., time-locked to a crucial position within the stimulus) before one can reasonably compare them. Finally, using short lists instead of clausal stimuli is advantageous in testing infants (Study III), since shorter and less complex stimuli are easier to access for a population with a short attention span and low processing capacities.

Moreover, comparable stimuli have already been used in behavioral studies on boundary perception (e.g., Aasland & Baum, 2003; Streeter, 1978; see above) testing the impact of prosodic boundary cues and cue combinations. Thus, the coordinate structures have proven suitable to unravel fine differentiations in boundary perception. In addition, production studies with English and German speakers found that the crucial prosodic boundary cues yielding the intended subgroupings are produced in coordinate structures. This has been shown for lists of short sentences (e.g., Féry & Truckenbrodt, 2005; Ladd, 1988) and for lists of names (e.g., Kentner & Féry, 2013; Petrone et al., in press; Wagner, 2005) as used in the current work.

On these grounds, the chosen material is considered most suitable to investigate PB perception using ERPs, both in infants and adults. However, care has to be taken when drawing inferences from different studies since the apparent methodological variety—also with regard to the stimulus material—may distort the picture and at least complicates direct comparisons.

4 Research questions and hypotheses

Despite the importance of prosodic phrasing for both language acquisition and comprehension, little is known about the procedural mechanisms underlying prosodic boundary perception. In order to specify factors that play a crucial role in perceiving a prosodic boundary as such, the present work investigates both contextual influences as well as the impact of specific prosodic cues and cue combinations on PB perception in German.

Study I addresses the role of the prosodic context and explores whether adults' perception of a prosodic boundary is affected by its position within an utterance. Specifically, the ERP experiment is aimed at answering the following questions:

- Is a CPS—indexing PB perception—elicited in response to PBs occurring in coordinated lists of names, irrespective of the boundary position within the coordinate structure?
- Does the previously processed prosodic context directly impact on the perception of a PB?

To answer these questions, Study I features systematic variation of the position of an utterance-internal PB and compares stimuli that contain an early PB, that is, a PB after a short phrase, with stimuli that contain a late PB which is thus preceded by a larger amount of prosodically structured speech material. Crucially, the local acoustic correlates of the prosodic boundaries do not differ across the two positions.

Given that coordinated lists of names have already been used to study PB perception behaviorally and that the relevant prosodic cues were evidenced in the production of comparable stimuli (see 3.2), I hypothesize that a CPS will be elicited in response to the PBs presented. Under the assumption that PB perception is solely based on the local occurrence of specific acoustic cues in the signal (e.g., Marcus & Hindle, 1990), I would expect a CPS in both conditions, with the latency of the component varying as a function of the boundary position (early vs. late). If, in contrast, PB perception depends on contextual factors such as the position of a prosodic boundary within an utterance, this should be evident in differences in the occurrence of the CPS between the two conditions that go beyond the presupposed latency difference. Hence, in Study I, the comparison between the ERPs in response to an early PB with those in response to the late PB is expected to shed light on potential processing differences that are due to the amount of previously processed speech material, that is, prosodic context information.

Study II examines the processing of specific prosodic boundary cues and cue combinations to specify their role in adult PB perception. In particular, two experiments collecting ERP and prosodic judgment data address the following questions:

- Do adult listeners perceive a PB in acoustically manipulated speech material that contains only a subset of the established prosodic cues?
- More precisely, is a CPS—indexing PB perception—elicited in response to PBs that are signaled by a) pitch change only, b) final lengthening only, or c) a combination of pitch change and final lengthening, but no pause cue?
- Do the ERP results match the prosodic judgments? That is to say, to what extent does the occurrence of the neurophysiological marker known to signal PB perception (i.e., the CPS) relate to the behavioral results?

Study II uses stimulus material that is systematically manipulated to determine the impact of the two single cues and the combined cue occurrence. With regard to the CPS, the reasoning goes as follows: If pitch change and final lengthening are sufficient to trigger boundary perception as sole cues or in combination, a CPS will be elicited for the respective experimental conditions as compared to a baseline condition containing no prosodic boundary cues. Based on previous behavioral and ERP research (see above), I hypothesize that a CPS occurs when pitch and final lengthening cues are jointly presented. If, in contrast, no CPS is elicited in response to the joint cue presentation, this would hint at an essential role of the third prosodic cue, namely the pause between two major prosodic phrases. Concerning the impact of the single prosodic cues, empirical evidence so far is inconsistent and, especially with regard to ERPs, very sparse. A CPS elicited in response to the respective condition would hence point to the single cue being sufficiently effective to trigger PB perception, while the absence of a CPS would indicate that pitch change and/or final lengthening are not effective as single cues. Crucially, if a CPS is elicited for more than one condition, CPS amplitude differences may shed light on an additive functioning and the relative importance of the prosodic cues. Moreover, with regard to the evaluation of on-line versus off-line methodological approaches, I expect both behavioral and ERP data to reflect PB perception concordantly. That is to say, I assume that the occurrence of the CPS will be reflected in the prosodic judgments given in response to the respective condition. Nevertheless, the two methodological approaches complement one another in that ERPs may provide additional insights into prosodic boundary cue processing when behavioral results are inconclusive (i.e., responses at floor or chance level).

Analogously, **Study III** examines the development of prosodic boundary cue perception in infants. To that end, six- and eight-month-olds are presented with stimuli containing either no boundary cues, only a pitch cue, or a combination of pitch change and final lengthening. Using ERPs, the study avoids the need for an overt (motor) response and seeks to clarify the following research questions:

- Is PB perception in six- and eight-month-old infants already reflected by a brain response similar to the adult CPS?
- Do German-learning infants perceive a PB either signaled by the co-occurrence of pitch change and final lengthening or by a pitch cue presented in isolation?
- Can the developmental path suggested by behavioral studies be traced on the neurophysiological level, that is, are there differences between the two age groups that can be interpreted in terms of an attunement in prosodic boundary processing?

To date, there is mixed evidence regarding the existence of an infant CPS (see Chapter 3.1.2). However, infants' ability to make use of prosodic boundary cues for segmentation has been demonstrated in numerous behavioral studies; it thus seems likely to be reflected in the infants' brain responses. In contrast to previous studies, using stimuli that do not contain a pause cue will allow to clearly tease apart a possible infant CPS from the occurrence of an obligatory onset component.

Under the assumption that previous behavioral results (Wellmann et al., 2012) will be reflected in the ERP data, I hypothesize that eight-month-old infants show a specific brain response indicating PB perception in response to stimuli that contain both pitch change and final lengthening. In contrast, the brain response to stimuli containing only the pitch cue will presumably not differ from that to stimuli without PB cues. Given that six-month-old infants only recognized PBs containing the pause cue (Wellmann et al., submitted), I assume that this age group will not show a specific brain response signaling PB perception in response to either of the conditions containing prosodic boundary cues. Hence, the developmental shift occurring behaviorally between six and eight months of age should be traceable in the ERP data. However, this assumption needs to be qualified by mentioning that a specific brain response may precede the corresponding behavioral response in the course of development (see, e.g., Männel & Friederici, 2008, for a methodological discussion) resulting in accordant but time-shifted observations at the behavioral and the neurophysiological level.

5 Summary of the major results

In **Study I**, adult participants were presented with coordinated lists of names that contained a PB either after the first (EARLY condition) or after the second name (LATE condition). Visual inspection of the ERP data suggested that a CPS is elicited in response to stimuli from the LATE condition, whereas no such positive shift occurred in response to the stimuli with a PB after the first name. Given that auditory ERPs are highly susceptible to acoustic changes and latency differences in the stimulus material (see, e.g., Steinhauer & Drury, 2012), three different approaches to a statistical data quantification were pursued in order to account for possible confounding stimuli differences: Two analyses assessed differences in the ERP data relative to stimulus onset, using different baselines and time windows, and a third analysis compared the ERP data time-locked to the boundary position. All analyses yielded statistically significant results that reflect differences in processing an early as compared to a late PB. Moreover, the very first broad analysis as well as the analysis relative to the boundary position statistically confirmed the occurrence of a CPS corresponding to the PB at the late boundary position. The third, most fine-grained analysis revealed a fronto-central distribution of the CPS effect. In contrast, no CPS occurred in response to the PB at the early boundary position. This positional effect—yielding a CPS in response to the LATE condition, but not in response to the EARLY condition—shows that PB perception depends on contextual factors such as the position of a boundary within an utterance. It follows that the prosodic cues, which were unequivocally present in both conditions, were processed in different ways. Given that the crucial difference caused by the manipulation of the boundary position was the length of the major prosodic phrase preceding the boundary, the results presumably reflect an immediate impact of this preceding prosodic context on PB perception. Despite this influence of prosodic context, the results of Study I clearly demonstrate that coordinate structures are suitable to investigate PB perception, as a CPS was reliably elicited for the PB after the second of three names.

Given the results of Study I, **Study II** focused on the late boundary position and examined the role of pitch change and final lengthening in German PB perception. In two experiments that combined ERP and behavioral measures (i.e., a prosodic judgment task), I investigated whether adult listeners perceive a PB in acoustically manipulated speech material that contained either no, one, or two prosodic boundary cues. Both the ERP and the behavioral results suggest that pitch change and final lengthening cues need to occur in combination to trigger PB perception: Concerning the neurophysiological findings, only the experimental conditions with a combined occurrence of both cues provoked a CPS response. Correspondingly, on the behavioral level these stimuli were predominantly judged as containing a PB. Pitch change and final lengthening presented as single boundary cues, in contrast, did not elicit a CPS nor did they shift the listeners' judgments towards perceiving a PB. This pattern of results was obtained regardless of the two types of pitch cue employed in the experiments (i.e., local vs. global pitch manipulation, see 8.2.2). Hence, behavioral and ERP data from Study II consistently show that German adult PB perception is evoked by the combination of pitch change and final lengthening, but not by pitch change or final lengthening alone.

Study III was aimed at investigating the development of prosodic cue perception in German-learning infants. Based on the findings from Study II, six- and eight-month-olds were—due to the lower attention span—presented with a subset of the established experimental conditions of Study II, that is, with stimuli containing either no boundary cues, only a pitch cue, or a combination of pitch change and final lengthening. For both age groups, the ERP results featured a positive deflection in response to the latter condition. Given that the statistical analyses revealed significant differences compared to the control condition without boundary cues, this suggests the presence of an infant CPS signaling PB perception provoked by co-occurring pitch change and final lengthening cues. In contrast, the ERP wave forms elicited in response to stimuli containing only the pitch cue resembled the ERPs obtained for the control condition; correspondingly, the two conditions did not significantly differ from each other. Overall, the findings of Study III indicate that both six- and eight-month-old infants perceive a PB when it is signaled by a co-occurrence of the two examined cues. Accordingly, PB perception at these ages does not appear to require the presence of a pause cue. Further on, in line with the results for adults, a pitch cue presented in isolation is insufficient for PB perception in both age groups.

6 Conclusions

This dissertation encompasses three ERP studies investigating factors that potentially impact on PB perception: contextual influence and the role of specific boundary cues and cue combinations. The experimental findings shed light on the procedural mechanisms underlying prosodic boundary perception. In what follows, I will first summarize what can be inferred from the respective results and how the findings relate to the research questions formulated in Chapter 4. To conclude, I will illustrate how the current work contributes to a better understanding of the role of prosodic boundary perception in language comprehension and acquisition, also considering potential limitations and open questions for future research.

6.1 On the role of prosodic context in PB perception

With regard to the influence of prosodic context on PB perception in coordinate structures, Study I revealed that prosodic boundaries are not processed locally in the sense that a PB is perceived whenever the relevant prosodic boundary cues are present. Instead, PB perception—as indexed by the occurrence of a CPS—was shown to be affected by the preceding prosodic context in that a CPS was only elicited in response to a late PB, but not in response to an early PB following a very short phrase.

To begin with, it should be noted that the stimuli in previous ERP studies on PB perception usually consisted of long sentences (e.g., Bögels et al., 2010; Männel & Friederici, 2009; Pannekamp et al., 2005; Steinhauer et al., 1999). Study I is the first work demonstrating that the CPS can be elicited in response to prosodic boundaries in short, non-sentential sequences, such as coordinated lists of names. In line with behavioral studies using this kind of coordinated structures to investigate prosodic phrasing in production and perception (e.g., Kentner & Féry, 2013; Lehiste, 1973; Wagner, 2005, 2010), this result of Study I provides further evidence for the CPS as an indication of PB perception.

Crucially, the observed pattern of results points to a direct contextual influence on PB perception: While the CPS indexes PB perception in the LATE condition, the absence of the CPS in response to a PB occurring at the early boundary position strongly suggests that the encountered acoustic boundary correlates—a pitch rise, final lengthening, and a pause—were not processed as PB cues when they followed a very short phrase early in the utterance. This processing difference suggests that at the early position, PB perception was not warranted by the preceding prosodic context. Two possible accounts explain the results either in the light of relative boundary processing or from a cognitive resource viewpoint (see also 7.4):

First, the acoustic correlates may not have been processed as PB cues because there was not enough previous prosodic information available serving as a benchmark to evaluate the acoustic changes as PB cues. This reasoning is in line with behavioral studies highlighting the importance of the relative strength of a PB compared to other nearby boundaries (e.g., Carlson, Clifton, & Frazier, 2001; Clifton et al., 2002; Wagner, 2005, 2010). Such a benchmark (i.e., in terms of a preceding weaker boundary) was not available within the first short constituent. Moreover, not only previous boundaries, but also contextual prosodic information in general may serve as a reference system for boundary perception. Due to the shortness of the first constituent, no such or not enough previous context (e.g., information on segmental duration and pitch variation) was available to interpret the prosodic boundary cues as such during on-line processing.

Second, the absence of the CPS may result from an unnecessity for chunking at the early boundary position. As outlined above, prosodic boundaries enable the listener to chunk the incoming auditory signal into larger units and may thus help to reduce processing costs. In the EARLY condition, the boundary cues are encountered when listeners have only perceived a minor part of the utterance (in fact, only two syllables). Therefore, there may simply be no need to chunk this word into a larger (prosodic) unit. Notably, it is likely that the cognitive process underlying prosodic chunking is reflected in CPS occurrence. Therefore, the unnecessity to chunk information leads to the absence of the CPS at this early position.

Indeed, both the notion of unnecessary chunking as well as the suggested account of lacking prosodic context maintain the core assumption that the CPS is not only mirroring lower level perceptual processes such as the encounter or detection of prosodic cues. Rather, it has a linguistic or cognitive relevance signaling the use of prosodic boundary information during on-line processing.

In sum, the results of Study I are in line with a non-local account of boundary processing (e.g., Clifton et al., 2002) assuming that prosodic boundaries are processed relative to previous prosodic information. Importantly, implications from previous behavioral results are extended by demonstrating that the impact of global prosodic context is not restricted to subsequent processes of sentence interpretation, but that it immediately influences the perception of a prosodic boundary.

6.2 Pitch change, final lengthening, and pause cues in adult and infant PB perception

The role of specific boundary cues and cue combinations in adult and infant boundary perception was investigated in Study II and III. In particular, four ERP experiments explored the impact of pitch change and final lengthening presented as single boundary cues and in combination, respectively, allowing for conclusions to be drawn concerning also the third main PB cue, namely, a pause. As for the infants, testing two different age groups aimed at exploring the developmental path of boundary cue perception in the course of language acquisition.

Concerning the contribution of specific prosodic cues to PB perception in adults, Study II revealed that listeners perceive a PB in acoustically manipulated speech material that contain a combination of pitch change and final lengthening, but no pause cue. Further, pitch change and final lengthening need to occur in combination, because the single cues were found to be insufficient to trigger PB perception. This pattern of results was obtained both at the behavioral and at the neurophysiological level, confirming that CPS occurrence closely relates to listeners' judgment of having perceived a PB: Only the experimental conditions containing a combination of pitch change and final lengthening gave rise to the occurrence of a CPS and, correspondingly, significantly shifted the listeners' judgments towards perceiving a PB.

With regard to the nonattendant pause cue, Study II confirms that—albeit being an especially salient cue—it is not a necessary cue to PB perception. This finding is in line with previous ERP studies (Männel & Friederici, 2009; Steinhauer et al., 1999) showing that a CPS is elicited even if the pause was cut out of the stimulus material.

Further, it matches a corpus study on the occurrence and distribution of prosodic phrase boundary cues in German spontaneous speech (Peters et al., 2005) which found pause to be optional and rather infrequent at turn-internal PBs.

As for the role of the single cues presented, the results seem to conflict with earlier behavioral studies postulating an impact of pitch change and final lengthening on PB perception, albeit being presented as single cues (e.g., Streeter, 1978; Yang et al., 2014). Crucially, this conclusion was drawn from the observation of a higher tendency to accept an alternative phrasing. Also in the current study, listeners showed significantly larger proportions of trials judged as containing a PB for the single cue conditions in comparison to the baseline condition without boundary cues. Thus, the current judgment results actually match previous behavioral findings in that the acoustic manipulations affected the judgment even when only one cue was manipulated, that is, presented in the stimulus material.

However, labeling the single prosodic cues tested here as "effective" in the sense that they may be sufficient to trigger PB perception would be misleading, since the larger proportion of judgments observed for the single cue conditions are still below or at chance level; notwithstanding that the acoustic differences between experimental conditions affect the participants' response behavior (e.g., resulting in a higher degree of uncertainty or guessing). In contrast, only the conditions containing both pitch change and final lengthening yielded results that were significantly above the chance level, reflecting consistent or at least predominant PB perception for these conditions. The interpretation that only the combined occurrence of pitch and lengthening is sufficient to trigger PB perception is also strongly supported by the neurophysiological data, that is, the absence of a CPS for the respective single cue conditions. Hence, combining behavioral and neurophysiological measures suggests a methodological refinement in the sense that a significant increase in the prosodic judgments towards perceiving a boundary or an alternative prosodic phrasing must not necessarily be reflecting consistent PB perception.

Moreover, differences in the experimental designs (i.e., using categorical vs. gradient boundary cue manipulation) and in the magnitude of the boundary cues provided may explain diverging results in previous studies. The current work used categorical cue manipulations (i.e., a cue was either absent or present) that were based on values found in the natural productions of the particular speaker. The aspired benefit of this design is to create stimuli that contain PBs that are as natural as possible while being able to control the impact of particular prosodic cues. Admittedly, the experimental outcome may depend on exactly such choices regarding the experimental design (see, e.g., Aasland and Baum, 2003; Petrone et al., in press, for stronger effects of single prosodic cues obtained using a gradient paradigm).

Finally, Study III examined the development of prosodic boundary cue processing within the second half of the first year of life. The ERP results suggest that PB perception in six- and eight-month-old infants is already reflected by a brain response similar to the adult CPS. Clearly, the observed effects cannot be attributed to low-level processes triggered by newly incoming acoustic material after silence. This explanation has previously been used by Männel and Friederici (2009) to account for a CPS-like deflection in eight-month-old infants' brain response to fully marked PBs (Pannekamp et al., 2006). Instead, the results of Study III support the idea of a CPS-like infant ERP component as proposed by Pannekamp et al. (2006): In eight-month-old infants, this brain response was elicited in response to stimuli containing a combination of pitch change and final lengthening but no pause cue. Thus, the ability to perceive a PB marked by this subset of prosodic cues, which has previously been demonstrated in behavioral studies on English and German (Seidl, 2007; Wellmann et al., 2012), is also reflected on the neurophysiological level. Accordingly, no CPS was elicited in response to stimuli containing only one of the boundary cues, that is, the pitch change cue. Hence, the infant CPS reflects the sensitivity to prosodic boundary information exploited in language acquisition. Presumably, it should also mirror the attunement of prosodic cue processing abilities related to the ambient language.

Surprisingly, however, six-month-olds showed the same pattern of results, namely a CPS in response to the combination of pitch change and final lengthening, but not in response to pitch change as a sole PB cue. This finding deviates from the hypotheses (see Chapter 4) based on behavioral results indicating that six-month-old German learners only recognize PBs that involve a pause cue ((Wellmann et al., submitted) Thus, the developmental shift observed behaviorally between six and eight months of age (Wellmann et al., 2012; Wellmann et al., submitted) is not reflected in the present ERP results. As explicated in Chapter 9.4, this discrepancy presumably originates from different cognitive demands associated with behavioral versus neurophysiological test procedures. Thus, a certain acquisition step may first be evident in ERPs, and, more indirectly assessed, only later observable in a behavioral response. For instance, ERPs indicate the recognition of the ambient language's dominant stress pattern in four-month-old German learners (Friederici, Friedrich, & Christophe, 2007) while the corresponding behavioral preference is evident only at six months (Höhle, Bijeljac-Babic, Herold, Weissenborn, & Nazzi, 2009). Therefore, the same pattern of results (i.e., a younger age group relying heavily on the pause cue while older infants perceive a PB on the basis of pitch change and final lengthening only) potentially occurs when comparing ERPs of even younger, for instance four-month-old infants with the data obtained for the six-month-olds.

Although the developmental shift—presumably grounded in a decline of reliance on the pause cue—is not apparent in the present ERP results, they still shed light on young learners' ability to make use of rather subtle, gradient prosodic cues. While there is much evidence for infants' particular sensitivity to pauses (e.g., Hirsh-Pasek et al., 1987; Morrongiello & Trehub, 1987; Schmitz, 2008), the current finding that German-learning infants consistently perceive a PB despite the lack of a pause hints at an early adult-like usage of relational prosodic cues, seeing that the pattern of results also resembles the adult results obtained in Study II.

6.3 General conclusion and future directions

This dissertation offers new scientific knowledge that highlights the role of prosodic context and evaluates the prosodic cues under investigation. Moreover, it provides methodological implications concerning research on the integration of prosodic information both in language comprehension and acquisition.

Regarding the methods applied, this dissertation reinforces previous studies on the functional significance of the CPS. In several respects, it yields evidence that the occurrence of a CPS does not merely reflect a brain response to acoustic changes (which in turn may indicate a PB). Instead, the CPS mirrors the integration of available prosodic boundary information into the parsing process. In other words, it signals the *use* of prosodic boundary cues for sentence comprehension. Hence, the CPS constitutes an excellent tool to investigate the cognitive mechanisms underlying prosody perception and the interplay of prosodic and syntactic as well as semantic processing. This insight is strengthened by the correspondence of ERP and behavioral data in Study II. Importantly, the present work is the first to show that this holds true also for rather short, non-sentential stimuli like coordinate structures. Therefore, future studies may also benefit from using this kind of consolidated and well-controlled stimulus material.

Moreover, the present work expands the potential scope of ERP studies focusing on the CPS, as it supports the notion of an infant CPS that reflects early linguistic processing of prosodic boundary cues. Making use of the infant CPS may allow future research to study boundary perception and the integration of prosodic information in a population that can otherwise (e.g., using behavioral measures) be assessed only very indirectly. However, given that previous studies came to contradictory conclusions with regard to the existence of an infant CPS, future work should also substantiate the concept of an infant ERP component signaling the use of prosodic information in language acquisition.

Beyond the methodological implications, the present work sheds light on the processing mechanisms underlying prosodic boundary perception and, more generally, the integration of prosodic information. First, it affirms previous work demonstrating the immediate integration of prosodic information as signaled by the CPS. Second and most notably, the results of Study I suggest an immediate impact of the preceding prosodic context on PB perception. In particular, they reveal that the supposed influence of global prosodic context exerts during primary boundary processing, and not or not only during subsequent sentence interpretation.

Hence, the findings point to a crucial role of prosodic context, one of two factors affecting PB perception that were investigated here. However, because phrase length and the presence of an additional boundary were both varied in the experimental manipulation (i.e., in the LATE condition the critical PB occurred after a longer phrase, but was also preceded by an additional weaker boundary), future research is needed to disentangle which specific characteristics of the prosodic context impact on PB perception. Hence, it may well be that the strength of a boundary is not only defined relative to other nearby boundaries (as proposed by theories assuming *relative boundary strength*, see, e.g., Snedeker & Casserly, 2010; Wagner & Crivellaro, 2010), but also or rather with respect to preceding prosodic context which provides information, for instance, about segmental durations and pitch variation also at non-boundary positions. Notably, the current work suggests that in excess of a mere contextual influence, the perception of a PB needs to be warranted by the preceding context. Whether this mechanism primarily originates from perceived prosodic benchmark information, or whether a need for chunking affects PB perception—determined, for instance, by cognitive resource limitations or linguistic expectation (see, e.g., Brown, Salverda, Dilley, & Tanenhaus, 2011)—should be a matter for future research.

Future research on boundary processing elaborating the idea of relative boundary strength will also relate to the second factor investigated in this dissertation, namely the role of specific prosodic boundary cues and cue combinations. In adult PB perception as well as in acquisition, the prosodic boundary cues *pitch change* and *final lengthening* seem to be highly interrelated. Presented in combination, they consistently trigger PB perception, while single cues are insufficient.

Presumably, this outcome relates at least partly to remarkable dependencies in the perception of pitch and duration: Clearly, pitch change information has to unfold over time to be perceivable. Beyond that, the pitch contour has also been found to affect the perception of (vowel, syllable, or non-speech sound) duration across different languages (see, e.g., Cumming, 2011; Lehiste, 1975; Pisoni, 1976; Yu, 2010; but also Lehnert-LeHouillier, 2007 for a failure to replicate the findings for German, Thai, and Spanish). Specifically, participants perceived stimuli as longer that had a dynamic fundamental frequency (as, e.g., in a pitch rise) compared to stimuli with a static fundamental frequency (i.e., without pitch change). Moreover, Brugos and Barnes (2012, 2014a) showed that pitch information distorts perceived duration not only for simple tones, but also in the perception of speech material. Accordingly, Cumming (2010) found tonal and durational cues to be interdependent in the perception of speech rhythm.

However, Brugos and Barnes (2014b) employed a phrasal grouping task and revealed that although pitch change modulates perceived duration, the impact of pitch change on perceived *grouping* goes clearly beyond effects that could be attributed to distortions of perceived duration. Therefore, the authors conclude that listeners integrate both pitch and durational cues to render prosodic grouping judgments.

Still, future research may revise the concept of cue weighting in PB perception, acknowledging that prosodic cue size is not sufficiently characterized in terms of acoustic-phonetic measures. Instead of trying to artificially tease apart the contribution of the acoustic correlates presented in isolation, the cue trading relationship, that is, the interplay between two different cues (i.e., pitch change and final lengthening; see, e.g, Beach, 1991; Cumming, 2010) at one critical boundary position or between the same acoustic correlates across different neighboring boundaries, is worth a closer look. Moreover, if prosodic boundaries are processed and evaluated relative to the preceding prosodic context, it clearly follows that the impact of a specific prosodic cue is also not primarily defined by the absolute size of a cue, but rather by its relative importance in a given context.

Another aspect that should be taken into account in future research on boundary cue processing is the role of interspeaker-variability. Speakers show a large variation in their use of pauses, lengthening, and pitch contour to signal PB boundaries (e.g., Peppe, Maxim, & Wells, 2000; Petrone et al., in press), which may in turn affect boundary perception. Although the present work did not aim to cover this issue, it should at least be mentioned that the conventional concept of cue weighting neglects interspeaker-variability in its attempt to determine the relative importance of prosodic boundary cues solely according to the prosodic characteristics of the ambient language.

Moreover, although the current findings have been compared to those of previous behavioral studies (see above), this does not allow for conclusions to be drawn on cross-linguistic variation, given the diverging experimental designs and measures. Rather, the current work may lay the foundation for prospective cross-linguistic research on boundary processing, as it focuses on the particular contribution of the three boundary cues to German PB perception.

While the perceptual interplay of pitch change and final lengthening is not yet well-understood, the decisive role of the combination of the two relational cues in both language comprehension and acquisition becomes obvious in the light of the current results. The pause—albeit being a categorical and thus more salient cue—has been shown to be an optional, non-essential cue, confirming previous ERP studies on German (Männel & Friederici, 2009; Steinhauer et al., 1999) that observed PB perception despite the absence of a pause cue.

One of the research aims was to tease apart ERP effects due to boundary perception from those reflecting low-level processing of signal re-occurrence after an interval of silence, especially concerning infant PB perception. For this reason, the present studies on the specific contribution of boundary cues involved only stimuli without a pause following the PB. Therefore, neither additional cue combinations (e.g, pause and final lengthening) nor pause as a single cue could be examined in the current paradigm. This clearly limits conclusions to be drawn on a possible cue weighting or the relative importance of pause as compared to pitch change and final lengthening. Nonetheless, the questions arises as to what accounts for the dispensability of the pause despite its perceptual salience.

One reason for this may be found in production data, relating the importance of prosodic cues in perception to their occurrence as a boundary cue in the ambient language. For instance, a corpus study of German spontaneous speech (Peters et al., 2005) found pause to be optional and rather inconsistent at turn-internal PBs. Alternatively, an acoustic correlate may constitute a weaker prosodic boundary cue when its occurrence is potentially ambiguous in the sense that it may also fulfill other tasks than prosodic phrasing. For instance, silent pauses occur as a reflection of speech planning processes (e.g., Goldman-Eisler, 1968) and are frequently found in hesitation disfluencies or preceding a disfluency repair (Nakatani & Hirschberg, 1994). Therefore, pauses presumably constitute a rather unreliable PB cue, at least with respect to turn-internal boundaries. Nonetheless, infants are particularly sensitive to pauses (e.g., Hirsh-Pasek et al., 1987; Morrongiello & Trehub, 1987; Schmitz, 2008) and use them for segmentation, presumably exploiting their categorical nature and acoustic salience. Given the outlined inconsistency of pauses as PB cue, the learner needs to enhance this segmentation strategy by paying attention to other, relational and thus more subtle boundary cues, namely pitch change and final lengthening. Accordingly, behavioral studies (for German: Wellmann et al., 2012; Wellmann et al, submitted) came to the conclusion that infants' PB perception develops from rather basic acoustic detection, which heavily relies on the pause cue, to a more sophisticated linguistic processing of prosodic boundary cues. The current finding that German learning six- and eight-month-old infants show a CPS reflecting PB perception despite the lack of a pause hints at an early overcoming of the exceeding reliance on pause occurrences. In lieu thereof, it indicates an early adult-like usage and linguistic processing of the relational prosodic cues signaling a major prosodic boundary.

II ORIGINAL JOURNAL ARTICLES

7 Brain response to prosodic boundary cues depends on boundary position[1]

Abstract

Prosodic information is crucial for spoken language comprehension and especially for syntactic parsing, because prosodic cues guide the hearer's syntactic analysis. The time course and mechanisms of this interplay of prosody and syntax are not yet well understood. In particular, there is an ongoing debate whether local prosodic cues are taken into account automatically or whether they are processed in relation to the global prosodic context in which they appear. The present study explores whether the perception of a prosodic boundary is affected by its position within an utterance. In an event-related potential (ERP) study we tested if the brain response evoked by the prosodic boundary differs when the boundary occurs early in a list of three names connected by conjunctions (i.e., after the first name) as compared to later in the utterance (i.e., after the second name). A *Closure Positive Shift* (CPS)—marking the processing of a prosodic phrase boundary—was elicited for stimuli with a late boundary, but not for stimuli with an early boundary. This result is further evidence for an immediate integration of prosodic information into the parsing of an utterance. In addition, it shows that the processing of prosodic boundary cues depends on the previously processed information from the preceding prosodic context.

---

[1] This chapter has been published as:
Holzgrefe, J., Wellmann, C., Petrone, C., Truckenbrodt, H., Höhle, B., & Wartenburger, I. (2013). Brain response to prosodic boundary cues depends on boundary position. *Frontiers in Psychology, 4*(421).

7.1 Introduction

Listeners' comprehension of spoken language is guided by prosodic information provided in the uttered speech stream. Prosodic characteristics such as the distribution of pauses or changes in the fundamental frequency have an important structuring function and thus help the listener to understand the speaker's intention. Utterances are chunked into prosodic phrases of different strengths (e.g., Nespor & Vogel, 1986; Selkirk, 1984, 2011; Truckenbrodt, 2007a), which helps to convey the correct meaning of a string of words. The boundaries of major prosodic phrases, so-called Intonational Phrases (Beckman & Pierrehumbert, 1986), are mainly signaled by three prosodic cues: a pitch change (i.e., a pitch rise or pitch fall indicates the presence of a boundary tone; this is usually followed by a pitch reset in the following phrase), final lengthening (i.e., an increase in the duration of the segments immediately preceding the boundary), and a pause (i.e., an interval of silence) between two phrases (see Peters et al., 2005 for German). Intonational phrase boundaries (IPBs) typically fall together with syntactic boundaries (Downing, 1970; Selkirk, 2005; for German: Truckenbrodt, 2005). For this reason, the perception of an IPB can be an important guide to the syntactic structure of spoken language; it is thus of special interest in psycholinguistic research in the attempt to bring to light how prosodic information is processed and how it contributes to sentence comprehension.

Numerous behavioral studies (see, amongst others, Carlson et al., 2001; Kjelgaard & Speer, 1999; Price et al., 1991; Schafer, 1997) have demonstrated an influence of prosodic boundary processing on syntactic analysis. Prosodic information is not present only at a local boundary, but instead unfolds throughout an utterance. Thus, the question arises of whether it is local boundary cues or rather prosodic information distributed across larger domains that has the primary influence on structural decisions during sentence processing. Proponents of the former view (e.g., Marcus & Hindle, 1990) have suggested that the processing of a prosodic boundary as a clue to syntactic structure is guided by the prosodic cues that occur in the direct vicinity of the boundary, regardless of other (prosodic) information that may be available to the listener. Prosodic boundary cues are thus supposed to be processed locally and context-independently. Others (e.g., Clifton et al., 2002) have argued against a solely local interpretation of prosodic boundaries. Instead, they provide evidence that the prosodic context in which an IPB occurs plays a major role, as the listener determines on a contextual basis whether a prosodic boundary is relevant for syntactic parsing decisions or not. In their work on the effect of prosodic information on the resolution of syntactic ambiguities, Clifton et al. (2002) identified two aspects that are relevant in this regard: the occurrence and the strength of neighboring prosodic boundaries, and the length of the prosodic phrase (e.g., the number of words or syllables) that precedes or follows a prosodic boundary.

Clifton et al. (2002) found that listeners interpret a prosodic boundary relative to preceding boundaries or potential boundaries within the same utterance. For example, in sentences with attachment ambiguities like *Old men and women with very large houses* a preference for a high attachment of the modifier *with very large houses* was found more often when the boundary before the modifier had not been preceded by a boundary after *men*. Moreover, not the strength of the boundary per se had an effect on the attachment decisions but the strength of the boundary relative to the preceding one. Hence, the occurrence as well as the strength of a preceding boundary affected the perception of a subsequent boundary, as revealed by the listeners' parsing preferences. More recent evidence for this impact of relative prosodic boundary strength on the perception of prosodic boundaries comes from Snedeker and Casserly (2010), as well as Wagner and Crivellaro (2010).

Phrase length, that is, the amount of material processed within one constituent, has been shown to affect prosodic phrasing in speech production and in the processing of implicit prosody in silent reading (Fodor, 1998; Gee & Grosjean, 1983; Hwang & Schafer, 2009; Watson & Gibson, 2004). Regarding the perception of prosodic boundaries, Clifton et al. (2006) demonstrate that phrase length affects the comprehension of syntactically ambiguous sentence structures. They presented participants in two auditory questionnaire experiments with sentences as in (4) and (5) (examples taken from Clifton et al., 2006; bracketing indicates the two different structures that were conveyed by prosodic phrasing):

(4) a. (Pat) or (Jay and Lee) convinced the bank president to extend the mortgage.

b. (Pat or Jay) (and Lee) convinced the bank president to extend the mortgage.

(5) a. (Patricia Jones) or (Jacqueline Frazier and Letitia Connolly) convinced the bank president to extend the mortgage.

b. (Patricia Jones or Jacqueline Frazier) and (Letitia Connolly) convinced the bank president to extend the mortgage.

The authors found a clear effect of the prosodic phrasing on sentence interpretation: participants were more likely to interpret stimuli in an "(X) or (Y and Z)" fashion when the prosodic phrasing suggests this analysis (Examples 4a and 5a), while the "(X or Y) and (Z)" reading was favored for the stimuli with the correspondent prosodic phrasing (Examples 4b and 5b; see also Lehiste, 1973 for the effect of prosodic phrasing on the interpretation of this kind of stimuli). Crucially, the effect of prosody was significantly larger for stimuli with short constituents (4a and 4b) as compared to stimuli with long constituents (5a and 5b). Clifton and colleagues interpret this result by assuming that listeners treat the boundaries flanking short constituents as more informative for the syntactic analysis, because long constituents could also be flanked by a prosodic break to assure speech fluency.

These findings strongly suggest that globally distributed prosodic information is integrated into the processing of prosodic boundaries as markers of syntactic structure. However, based on the data so far it cannot be decided whether the global prosodic structure has a direct impact on the perception and processing of prosodic boundaries or whether the effects observed in the data by Clifton and colleagues occur later during the process of sentence interpretation. This shortcoming is due to the limitations that apply to off-line methods such as judgment and reaction time data. Here, on-line methods with a high temporal resolution, like event-related potentials (ERPs), are a useful tool to unravel the time course of a potential influence of the global prosodic structure on the perception of boundaries. Therefore, the present study uses ERPs to investigate the perception of prosodic boundary cues at different utterance positions, varying the phrase length and thereby the amount of contextual prosodic information given before an IPB occurs. To illustrate that ERPs are useful in addressing questions on the integration of prosodic information and, in particular, on the time course of prosodic phrase boundary processing, the following section briefly outlines previous ERP research on prosodic boundary processing.

Studies on sentence processing using ERPs demonstrate an early influence of prosodic information on the syntactic parsing process (see, e.g., Eckstein & Friederici, 2006). Considerable evidence for this stems from studies on the perception of prosodic boundaries. Steinhauer et al. (1999) conducted an ERP study in which they compared German sentences that contained either one or two IPBs. Brain responses to prosodic violations (here, a prosodic boundary inserted at a non-boundary position) showed that syntactic processing was misled by prosodic information at an early processing stage. Crucially, as a response to each IPB, the authors found a broadly distributed, large positive waveform. Because the ERP component coincides with the closure of major prosodic phrases, it has been termed closure positive shift (CPS). The CPS has been found to indicate the processing of prosodic boundaries in various languages: in German (e.g., Männel & Friederici, 2009; Pannekamp et al., 2005), English (e.g., Itzhak et al., 2010; Pauker et al., 2011; Steinhauer, Abada, Pauker, Itzhak, & Baum, 2010), Dutch (e.g., Bögels et al., 2010; Kerkhofs et al., 2007), Japanese (Wolff et al., 2008), Chinese (Li & Yang, 2009), and Korean (Hwang & Steinhauer, 2011, implicit prosody).

Pannekamp et al. (2005) presented participants with sentences comparable to the material used by Steinhauer et al. (1999). However, in addition to the natural condition, their stimulus material was systematically varied: experiments were carried out using jabberwocky sentences (stimuli without semantic content, but with appropriate use of functional morphemes), pseudo-word sentences (containing neither semantic nor syntactic information), and hummed speech (without segmental information). Although the scalp distribution varied for the conditions that provided less linguistic information, the CPS was elicited in all four conditions. This shows that the CPS component occurs independently of semantic, syntactic, or segmental information. Moreover, Steinhauer et al. (1999) as well as Männel and Friederici (2009) demonstrate that a pause between two intonational phrases is not necessary to elicit a CPS. Hence, the CPS is not a variant of early auditory evoked potentials that signals the detection of new auditory input (e.g., after a break) and does thus not reflect lower level acoustic processing.

Concerning the impact of contextual prosodic information on prosodic phrasing, Hwang and Steinhauer (2011) used the CPS to demonstrate an influence of phrase length on (implicit) prosodic phrasing (which had previously been shown in behavioral production studies, see above). In a silent reading study on Korean they found that only longer sentence-initial constituents elicited a CPS, while no effect was found for short subject noun phrases. Here, the CPS was considered to reflect the subvocal generation of an additional prosodic boundary, which was only triggered by long constituents.

To summarize, ERP studies support the notion of an early integration of prosodic information in general. Furthermore, they provide converging evidence that the CPS component reflects prosodic boundary processing, as the occurrence of the CPS depends on prosodic information in the speech input, that is, the prosodic cues that mark the closure of an IPB, but not on other linguistic information or on mere acoustic changes in the input. In addition, the CPS has been shown to be sensitive to contextual effects on prosodic phrasing in silent reading.

Based on these findings, the present study makes use of the CPS as an indicator of IPB processing in differing prosodic contexts. In contrast to the complex sentences used in previous studies, our study employed coordinated lists with different syntactic and semantic subgroupings of the elements. The production of prosodic boundary cues in such subgroupings is shown in Ladd (1988) and Féry and Truckenbrodt (2005)—where the lists are lists of sentences—and in Wagner (2005) and Kentner and Féry (2013), where the lists are lists of names as in our experiment. We varied the position of the utterance-internal IPB in our stimulus material to determine whether prosodic context affects the processing of boundary cues as reflected by the occurrence of the CPS. We compared stimuli that contained an early IPB, that is, after a short intonational phrase, with stimuli that contained a late IPB, that is, an IPB preceded by a larger amount of prosodically structured material. Crucially, the local acoustic markers of the prosodic boundaries did not differ across the two positions. The reasoning goes as follows: If an IPB is processed solely based on the local occurrence of specific acoustic cues in the signal, we would expect the positive deflection to occur in both conditions with the latency of the component varying as a function of the boundary position (early vs. late). If, in contrast, IPB processing is affected by the boundary position, we should see differences in the occurrence of the CPS between the two conditions that go well beyond the presupposed latency difference.

7.2 Methods

7.2.1 Participants

Eighteen students of the University of Potsdam (12 women, age range: 20-28 years, mean age: 24.0 years) participated after giving informed consent. They were native speakers of German with no reported hearing or neurological disorders. All participants were right-handed, as assessed by a German version of the Edinburgh Handedness Inventory (Oldfield, 1971), and received course credits or reimbursement for their participation.

7.2.2 Material

Each experimental item consisted in a list of three disyllabic, trochaic names that were connected by *oder* ('or') and *und* ('and'). There were two experimental conditions that differed with respect to the prosodic grouping of the three names: in (6a), the EARLY condition, an IPB—signaled by a pitch change, final lengthening, and a pause—occurred after the first name, while in (6b), the LATE condition, the IPB occurred after the second name of the list (the position of the IPB is indicated by a hash mark in the examples):

(6) a. EARLY condition: [Mona]$_{IP}$ # [oder Lena und Lola]$_{IP}$
  b. LATE condition: [Mona oder Lena]$_{IP}$ # [und Lola]$_{IP}$

Six German names (*Lola, Lena, Lilli, Manu, Mona,* and *Nina*) were used to construct six different lists of three names. Hence, not all possible combinations of names were used, but it was ensured that each name occurred once in the first, the second, and the utterance-final position. All names were composed of four sonorants to allow a thorough acoustic analysis of the experimental material (see below). The six different lists of names were recorded in both prosodic conditions. During the experiment, each of these items was presented ten times, yielding a total of 60 experimental items per condition.

The stimuli were recorded in an anechoic booth by a naïve female native speaker of German. To ensure that the speaker produced the name sequences with the intended prosodic structure (early vs. late IPB), she was provided with a written list of the stimuli in which the intended prosodic grouping was indicated by bracketing, that is, *(Mona) (oder Lena und Lola)* for the EARLY condition and *(Mona oder Lena) (und Lola)* for the LATE condition. Each stimulus was preceded by the same context question (*Wer kommt?* 'Who is coming?') read by the experimenter. The speaker was instructed to read the name triples in such a way that the experimenter (who could not see the speaker's text) was able to mark the indicated grouping by adding the brackets in her written version of the stimulus list.

An example of a typical minimal pair is displayed in Figure 7.1. The figure shows that in both EARLY and LATE conditions the IPB is signaled by the presence of a silent pause (marked by a hash mark in the segmental labeling tier), pitch change (instantiated as a pitch rise and a reset after the pause) and lengthening of the preboundary segment.

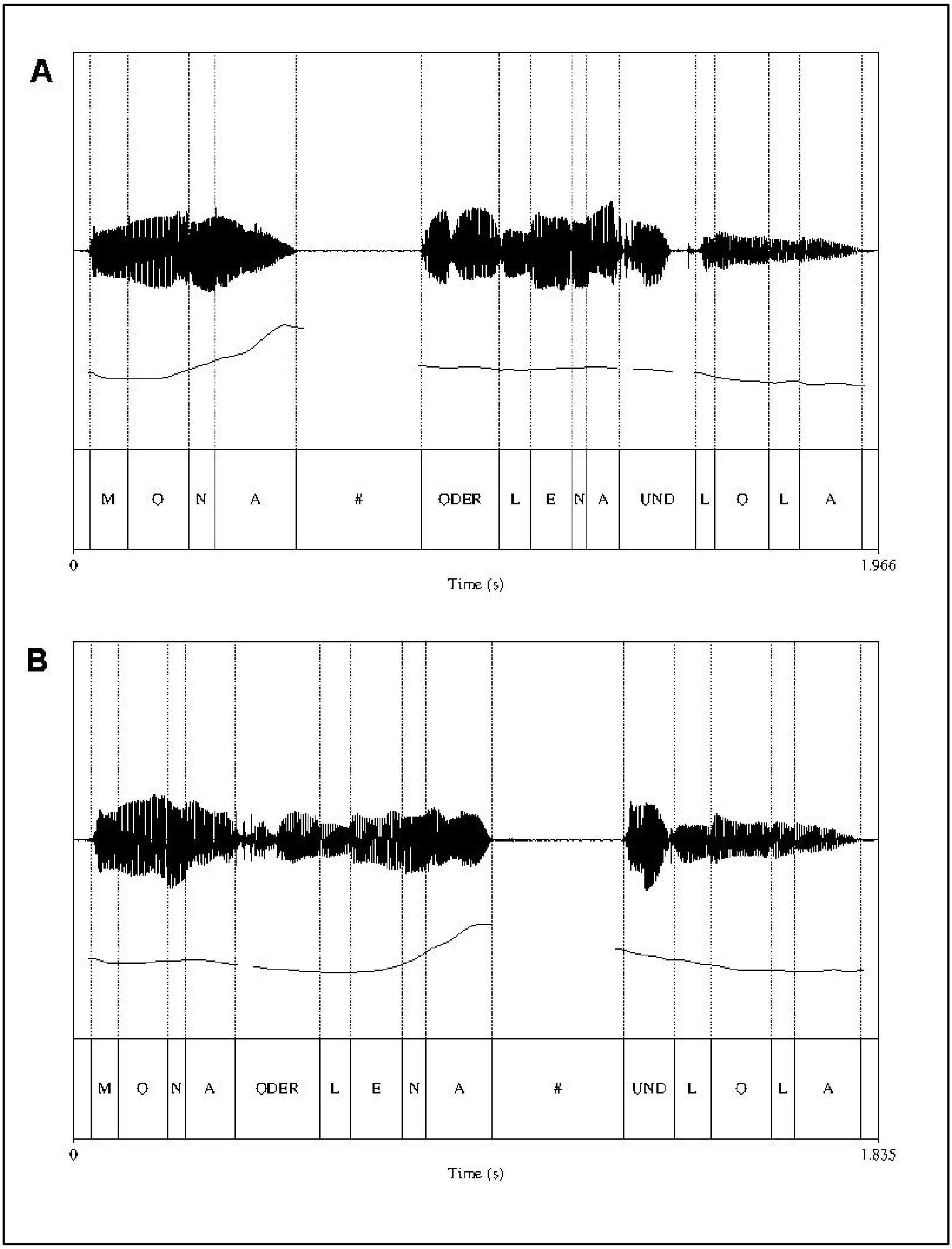

Figure 7.1 | Waveform, pitch track, and segmental labeling for the EARLY (panel A) and LATE (panel B) IPB conditions. Dotted lines mark the segmental boundaries. The silent pause after the IP boundary is indicated by a hash mark.

Acoustic analyses were carried out with Praat (Boersma & Weenink, 2010) to confirm that the relevant boundary cues—pitch change, final lengthening, and pause—were present and that the items from both conditions only differed in the critical respect, that is, the position of the prosodic boundary. An overview of the results is given in Table 7.1.

The durational properties—pause and final lengthening—were assessed by measuring the length of the final vowel of the first and the second name, as well as the length of a possible subsequent pause. In the EARLY condition, the final vowel duration of the first name was more than twice as long as on the second name and was followed by an extended pause, whereas no pause occurred after the second name. In the LATE condition, we observed the reversed pattern: the mean final vowel duration of the first name was shorter than the final vowel duration of the second name, which was again followed by a pause.

To assess the pitch change, we measured the preboundary pitch rise which occurred on the names that were potentially followed by an IPB. Therefore, the minimum of the fundamental frequency on the first sonorant of each first and second name was measured, as well as the maximum of the fundamental frequency on the final vowel (i.e., a high boundary tone). The difference of these values was used to calculate the pitch rise preceding the potential boundary position. In the EARLY condition, a major pitch rise occurred at the early boundary position: the pitch rise on the first name was almost five times as large as the slight rise measured on the second name (with no subsequent IPB). In the LATE condition, a comparably large pitch rise was observed on the second name, again in contrast to only a slight pitch rise on the first name.

Table 7.1 | Mean acoustic correlates of prosodic cues in the experimental stimuli.

Acoustic correlate	**EARLY condition**		**LATE condition**	
	First name	**Second name**	**First name**	**Second name**
Pitch rise in Hz (*SD*)	**144 (*21*)**	31 (*23*)	26 (*16*)	**151 (*21*)**
Maximum pitch in Hz (*SD*)	**350 (*21*)**	263 (*23*)	241 (*20*)	**340 (*19*)**
Final vowel duration in ms (*SD*)	**172 (*21*)**	84 (*14*)	113 (*13*)	**152 (*17*)**
Pause duration in ms (*SD*)	**297 (*29*)**	—	—	**268 (*20*)**

Numbers in bold represent measures from the IPB present in the respective condition.

The slight but perceivable pitch rise at the non-IPB positions (i.e., the second name in the EARLY condition and the first name in the LATE condition, see Table 7.1) hints at the presence of different tonal events at these positions. In particular, according to Truckenbrodt (2007b) it is attributed to the presence of a pitch accent and an edge tone of the accent domain. Such domain is the Accentual Phrase (AP), which is a prosodic constituent lower than the IP. Hence, each prosodic word in our material constitutes an AP, as illustrated in (7).

(7) a. EARLY condition: $[(\text{Mona})_{AP}]_{IP}$ # $[(\text{oder Lena})_{AP}$ $(\text{und Lola})_{AP}]_{IP}$

b. LATE condition: $[(\text{Mona})_{AP}$ $(\text{oder Lena})_{AP}]_{IP}$ # $[(\text{und Lola}_{AP})]_{IP}$

In sum, the acoustic analyses confirmed that the relevant IPB cues were present and did not differ in strength between conditions. There was only a positional difference: in the EARLY condition the crucial prosodic boundary cues—pitch change, final lengthening, and pause—were present at the end of the first name, while in the LATE condition they occurred at the end of the second name in the sequence. Taking the offset of the name before the IPB as indication of the IPB position, the positional difference amounts to ~500 ms: in the EARLY condition, the first name ends on average 488 ms (*SD* = 57 ms) after stimulus onset, while the second name in the LATE condition ends on average 992 ms (*SD* = 49 ms) after stimulus onset. Since latency differences—even if intended in the experimental design of stimulus material—play an important role in the interpretation of grand average ERPs, duration measures for the critical utterance parts in the experimental material are presented in Table 7.2. It becomes obvious that the stimuli systematically differ in critical word durations (i.e., due to final lengthening noun phrases are longer at IPB positions than at non-boundary positions) but not, for example, in total length. Moreover, latency differences occur between conditions, whereas duration measures within conditions are relatively homogenous.

Table 7.2 | Duration in ms for critical words, pauses, and utterance parts (before/after the pause), rounded to the nearest whole number, for each token employed in the EARLY Condition (EA1 to EA6) and in the LATE Condition (LA1 to LA6), respectively.

Experimental stimulus	**Duration for critical words in ms**				**Duration for utterance parts in ms**			
	NP1	Conj1	NP2	Conj2 +NP3	First part	Pause	Second part	Total
EARLY condition								
EA1 **[Mona] [oder Lena und Lola]**	545	189	287	639	545	306	1115	1966
EA2 **[Lena] [oder Lola und Mona]**	496	172	315	618	496	311	1104	1911
EA3 **[Lola] [oder Mona und Lena]**	509	160	274	547	509	273	981	1763
EA4 **[Nina] [oder Lilli und Manu]**	472	157	261	631	472	322	1048	1842
EA5 **[Lilli] [oder Manu und Nina]**	383	179	320	609	383	318	1108	1808
EA6 **[Manu] [oder Nina und Lilli]**	522	167	292	461	522	250	921	1693
Mean	**488**	**171**	**291**	**584**	**488**	**297**	**1046**	**1830**
SD	**57**	**12**	**23**	**68**	**57**	**29**	**80**	**99**
LATE condition								
LA1 **[Mona oder Lena] [und Lola]**	368	193	397	580	958	297	580	1835
LA2 **[Lena oder Lola] [und Mona]**	434	181	422	629	1037	259	629	1924
LA3 **[Lola oder Mona] [und Lena]**	382	228	402	568	1012	277	568	1857
LA4 **[Nina oder Lilli] [und Manu]**	362	196	352	611	910	269	611	1790
LA5 **[Lilli oder Manu] [und Nina]**	392	194	418	579	1004	267	579	1851
LA6 **[Manu oder Nina] [und Lilli]**	397	227	408	525	1031	237	525	1793
Mean	**389**	**203**	**400**	**582**	**992**	**268**	**582**	**1842**
SD	**26**	**20**	**25**	**36**	**49**	**20**	**36**	**49**

Note that for the EARLY condition duration of NP1 is identical to the first part of the utterance and for the LATE condition duration of Conj2+NP3 is identical to the second utterance part.

7.2.3 Procedure

The 120 experimental items were presented aurally (using E-A-RTONE 3A Insert Earphones, Aearo Technologies Auditory Systems, Indianapolis, USA) in a pseudo-randomized order with an inter-stimulus-interval of 4000 ms. The same sequence of names never occurred in consecutive trials and at most three consecutive trials belonged to the same condition. Participants were instructed to listen carefully and to avoid eye blinking and other body movements during stimulus presentation. To minimize eye movements, a fixation cross was displayed in the center of a monitor starting 1500 ms before stimulus onset until the end of the respective trial. The experiment lasted ~12 minutes.

The electroencephalogram (EEG) was continuously recorded from 30 cap-mounted active Ag/AgCl electrodes (Brain Products, Gilching, Germany) with a sampling rate of 1000 Hz. Electrodes were placed into the EEG cap at the following positions: Fp1/2, F7/8, F5/6, F3/4, Fz, FC3/4, FCz, T7/8, C3/4, C5/6, Cz, CP3/4, CPz, P7/8, P3/4, Pz, POz, O1/2. The electrooculogram (EOG) was recorded from electrodes placed above and below the right eye. Impedances were kept below 5 kΩ. The EEG recording was referenced on-line to the left mastoid and re-referenced off-line to linked mastoid electrodes.

7.2.4 Data analysis

The EEG data were analyzed using Brain Vision Analyzer (version 2.01; Brain Products, Gilching, Germany). A digital band pass filter ranging from 0.2 Hz to 70 Hz was applied to remove very slow drifts and muscle artifacts, and we also applied a 50 Hz notch filter. Epochs of 2200 ms, relative to stimulus onset, were extracted from the continuous EEG signal. Eye blinks and eye movements in the epochs were corrected by a computer algorithm (Gratton, Coles, & Donchin, 1983). All other artifacts were detected manually and contaminated segments were excluded from further analysis. The mean number of averaged trials per participant was 51.9 for the EARLY condition (SD = 5.4; 86.5 %) and 52.1 for the LATE condition (SD = 5.3; 86.8 %). The data of three additional participants were excluded from further analysis because the criterion of at least 40 artifact-free trials per condition (67%) was not met.

Two types of statistical analyses were performed that have been applied before to quantify CPS effects: first, we conducted analyses time-locked to the stimulus onset with a prestimulus baseline of 200 ms (see, e. g., ERP analyses in Männel & Friederici 2009, 2011; Pannekamp et al. 2005; Steinhauer et al., 1999), as well as adjusted to a baseline from 200 to 400 ms after stimulus onset, covering the early onset components. Second, additional analyses relative to potential boundary positions within the stimuli were conducted (see Bögels et al., 2010; Kerkhofs et al., 2007; Pauker et al., 2011, for comparable analyses), that is, time-locked to the offset of the first and the second name. In both cases, separate analyses were applied to lateral and midline electrodes. The following electrodes were used in the statistical analysis of lateral sites and were—by crossing the factors Region (anterior vs. central vs. posterior) and Hemisphere (left vs. right)—subdivided into six regions of interest: left anterior (F3, F7), right anterior (F4, F8), left central (FC3, C3), right central (FC4, C4), left posterior (CP3, P3) and right posterior (CP4, P4). In contrast, the separate analysis of the midline electrodes contained four levels of the factor electrode (Fz vs. FCz vs. Cz vs. Pz).

7.3 Results

7.3.1 Analyses relative to stimulus onset

Descriptive Results

The grand average ERP waves adjusted to a prestimulus baseline of 200 ms at the 16 electrodes used in the statistical analyses are illustrated in Figure 7.2. Additionally, voltage maps of differences waves based on all electrodes are shown in Figure 7.3, illustrating the scalp distribution of the amplitude difference between conditions.

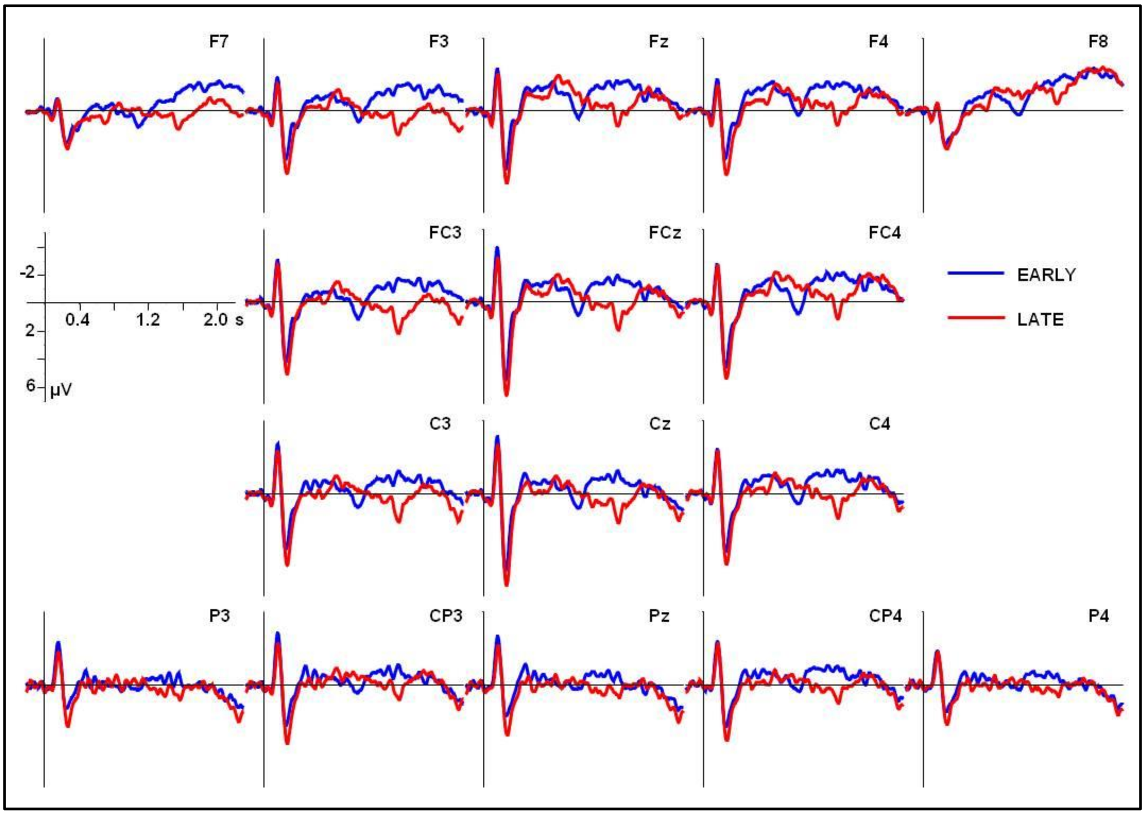

Figure 7.2 | Grand average ERPs adjusted to a prestimulus baseline of 200 ms for both conditions at the electrodes used in the statistical analyses. In all ERP figures an 8-Hz low-pass Butterworth zero-phase filter was applied off-line only for presentation purposes; all statistical analyses were performed on unfiltered data.

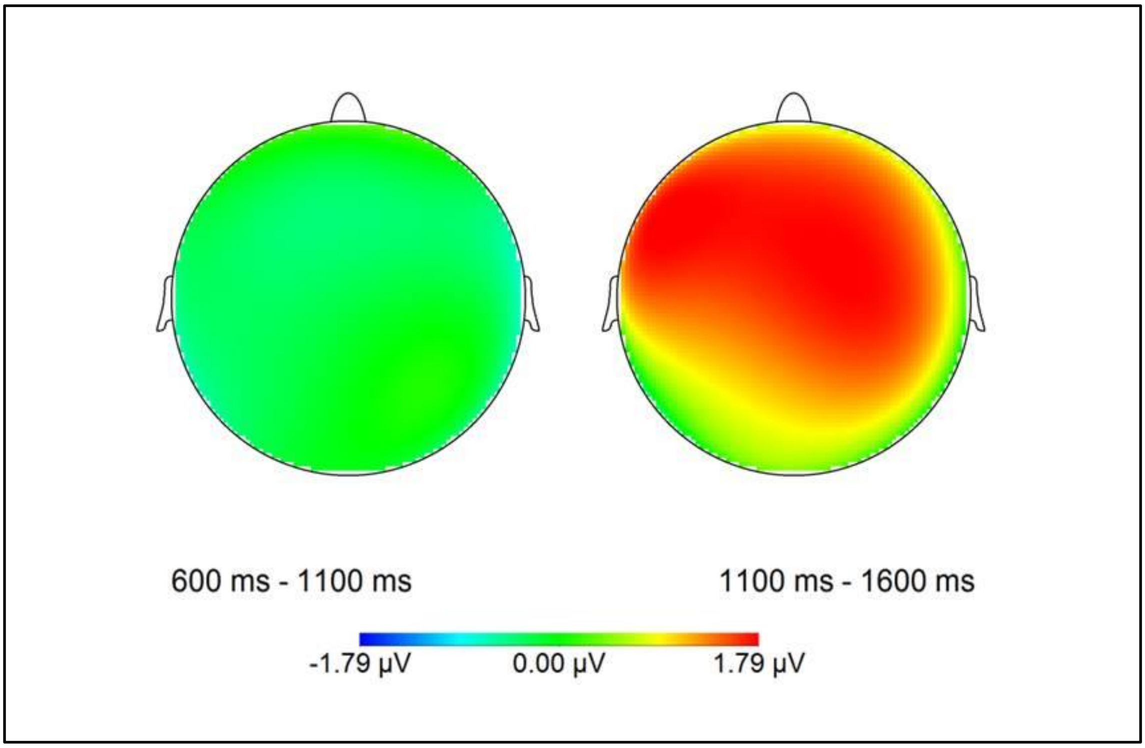

Figure 7.3 | Voltage maps of difference waves (adjusted to a prestimulus baseline of 200 ms) for the critical time windows used in the first statistical analysis relative to stimulus onset.

In Figure 7.4, grand average ERP responses at the representative Cz electrode are displayed. The dotted lines mark the mean pause interval in the two conditions, which lasted from ~500 to 800 ms after stimulus onset in the EARLY condition and from ~1000 to 1300 ms in the LATE condition. For both conditions, the obligatory N100-P200 complex (part of the auditory evoked potential, AEP; see, e.g., Picton, Hillyard, Krausz, & Galambos, 1974) is evoked in response to the stimulus onset from ~100 to 300 ms. Moreover, ERPs in both conditions display this obligatory components in response to the onset of the second part of the utterance after the pause. Here, the N100-P200 complex is less pronounced, presumably because it reflects a new onset within the utterance (as compared to utterance-initial) and because the pause duration slightly varies over stimuli (see Table 7.2). Still, a clear combination of a negative deflection followed by a positive peak can be found within the first 300 ms after pause offset—that is, around 1000 ms after stimulus onset in the EARLY condition and around 1500 ms in the LATE condition (to the right of the respective pause intervals indicated in Figure 7.4). To illustrate the onset components, the grand average ERP waves to each condition are depicted separately in Figure 7.5A and 7.5B.

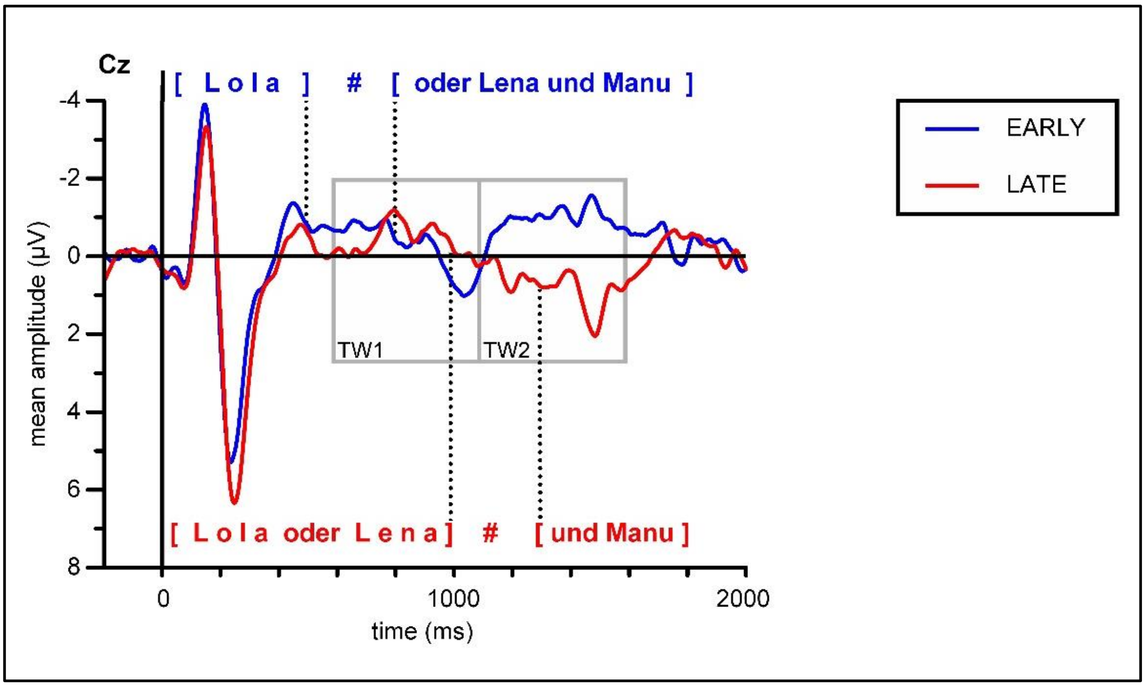

Figure 7.4 | Grand average ERPs for both conditions (adjusted to a prestimulus baseline of 200 ms) at electrode Cz. Grey boxes indicate the time windows used in the statistical analysis relative to stimulus onset. Dotted lines indicate the mean onset and offset of the pause at the IPB in the respective condition.

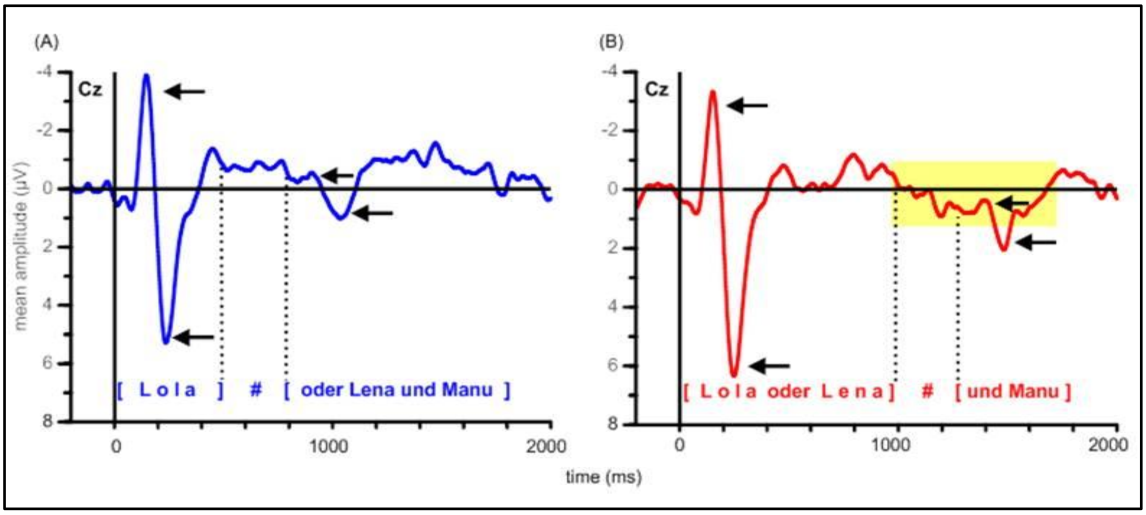

Figure 7.5 | Grand average ERPs (adjusted to a prestimulus baseline of 200 ms) at electrode Cz, depicted separately for (A) the EARLY and (B) the LATE condition. Arrows indicate the N100 and P200 components at the stimulus onset and at the onset of the second part of the utterance. Dotted lines delimit the mean pause intervals. In panel (B), the yellow rectangle indicates the time interval in which a positive shift can be observed, starting with pause onset and lasting for ~700 ms.

In addition to the obligatory components, a broad positive deflection can be observed for the LATE condition. It starts with the end of the first utterance part at around 1000 ms and lasts for ~700 ms (see Figure 7.4 and 7.5B). In the EARLY condition, a corresponding positive deflection that coincides with the offset of the first utterance part should start at around 500 ms after stimulus onset (see Table 7.2). As can be seen in Figure 7.4 and 7.5A, no such broad positivity is present in the EARLY condition. Hence, a positive shift coinciding with the IPB can only be observed for the LATE condition.

Statistical analyses relative to stimulus onset

For the statistical analysis relative to stimulus onset, epochs of 2000 ms were adjusted to a prestimulus baseline of 200 ms. Two consecutive time windows of 500 ms were defined in line with the possible occurrence of a CPS in response to the two experimental conditions. Remember that the IPB position in the EARLY condition, that is, the offset of the first name, was on average at 488 ms after stimulus onset, while in the LATE condition, the IPB occurred at the end of the second name, on average 992 ms after stimulus onset. Given these different IPB positions in the stimulus material, a CPS in response to the IPB in the EARLY condition should be revealed by statistical analyses of the first time window (TW1, 600 to 1100 ms after stimulus onset), while a CPS in response to the IPB in the LATE condition should lead to an effect of condition in the second time window (TW2, 1100 to 1600 ms).

A fully crossed repeated measures ANOVA was computed with the factors Time window (TW1 vs. TW2), Condition (EARLY vs. LATE boundary), Region (anterior, central, posterior), and Hemisphere (left vs. right); participants were entered as a random factor. The same analysis was conducted for the midline electrodes except that instead of the factors Region and Hemisphere only the factor Electrode (Fz vs. FCz vs. Cz. vs. Pz) was included. Subsequently, significant interactions involving the factor Condition were further analyzed using ANOVAs involving the respective factors. Only significant amplitude differences involving the factor Condition are reported. Where appropriate, a correction according to Greenhouse and Geisser (1959) was applied and reported as the corrected significance.

For lateral sites, the ANOVA including the factors Time window, Condition, Region, and Hemisphere revealed a statistically significant interaction of Time window x Condition [$F(1,17) = 13.25$, $p < .01$]. For midline electrodes, an ANOVA including the factors Time window, Condition, and Electrode revealed a significant interaction of Time window x Condition x Electrode [$F(3,51) = 4.31$, $p < .05$] and a significant interaction of Time window x Condition [$F(1,17) = 12.45$, $p < .01$].

To test the interaction with the factor Time window, subsequent statistical analyses were carried out on each time window separately. For both time windows a one-way ANOVA with the factor Condition was computed for lateral sites and a two-way ANOVA including the factors Condition and Electrode for the midline electrodes.

For the first time window (600 to 1100 ms), neither at lateral nor at midline sites was a significant main effect of Condition present, nor an interaction of Condition x Electrode at the midline electrodes, suggesting no differences between conditions at the early boundary position. For the second time window (1100 to 1600 ms), a statistically significant main effect of Condition was present for lateral electrode sites [$F(1,17) = 8.82$, $p < .01$] as well as for the midline electrodes [$F(1,17) = 6.72$, $p < .05$] with mean amplitudes in the LATE condition being more positive than in the EARLY condition.

In sum, this analysis relative to sentence onset indicated the occurrence of a broadly distributed CPS corresponding to the IPB at the LATE boundary position, whereas in response to the IPB at the EARLY boundary position, no positive shift occurred.

However, one reviewer suggested additional analyses with a baseline of 200 to 400 ms after stimulus onset instead of a prestimulus baseline to compensate for differences in the ERP wave forms occurring early after stimulus onset (see Figure 7.4). Figure 7.6 depicts the grand average ERPs for both conditions, adjusted to the 200–400 ms baseline. Moreover, slightly different time windows were proposed to quantify the CPS effects, with TW1 ranging from 700 to 1150 ms and TW2 ranging from 1150 to 1600 ms after stimulus onset. Paralleling the initial analysis, two fully crossed repeated measures ANOVA were computed separately over lateral and midline electrodes, including the factors Time window (TW1 vs. TW2) and Condition (EARLY vs. LATE boundary); participants were entered as a random factor. The statistical analysis employing the new baseline and slightly different time windows revealed a statistically significant interaction of Time window x Condition for lateral [$F(1,17) = 19.79$, $p < .001$] as well as for midline electrodes [$F(1,17) = 17.64$, $p < .001$]. Subsequent statistical analyses testing the interaction with the factor Time window were carried out on each time window separately. In contrast to the previous analysis using a prestimulus baseline, for the first time window (700 to 1150 ms) a significant main effect of Condition was present at lateral [$F(1,17) = 14.86$, $p < .01$] and midline electrodes [$F(1,17) = 10.82592$, $p < .01$]. For the second time window (1150 to 1600 ms) a statistically significant main effect of Condition was only present for lateral electrode sites [$F(1,17) = 4.82$, $p < .05$] but failed to reach significance for midline sites [$F(1,17) = 2.34$, $p = .1442325$]. Hence, the additional analysis using a different baseline period and differing time windows does not support the findings from the initial analysis. Instead, it points at differences between conditions in both time windows with an even more pronounced effect in the early time window. Therefore, the statistical data analyses relative to sentence onset lead to inconclusive results depending on the choice of baseline period and/or time window.

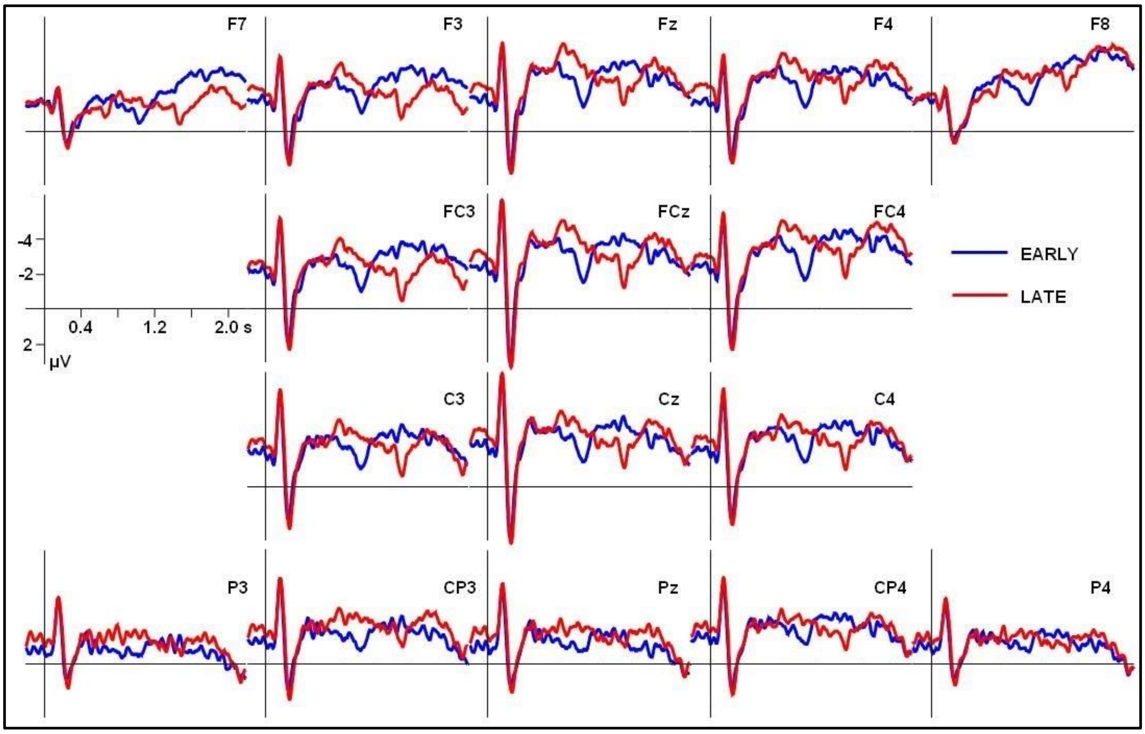

Figure 7.6 | Grand average ERPs for both conditions adjusted to a baseline from 200 to 400 ms after stimulus onset, covering stimulus initial onset components.

This highlights the importance of a thorough ERP data quantification especially regarding the choice of baseline and time windows employed (see also Steinhauer & Drury, 2012). Special care has to be taken in the interpretation of auditory ERPs since they are susceptible to acoustic changes and latency differences in the stimulus material. First, the investigation of prosodic boundary processing virtually always comes with critical latency differences in the stimulus material, because noun phrases before an IPB are longer than noun phrases at non-boundary positions (due to final lengthening, see material section above). Time-locking the ERPs to the boundary position, for example to the offset of the noun phrase followed by an IPB, allows to compensate for these latency differences between conditions (see Bögels et al., 2010; Kerkhofs et al., 2007, for comparable analyses). This may especially be necessary if subsequent boundary positions occur within one stimulus as in the present study. Second, it is necessary to disentangle the CPS from the P200 component (see, e.g., Picton et al., 1974) in response to the speech onset after the IPB (see Kerkhofs et al., 2007; Männel & Friederici, 2009; Pauker et al., 2011; Steinhauer, 2003). The previously described analyses did not meet this requirement, because subsequent P200 components occur within the time windows chosen to quantify possible CPS effects, as can be seen in Figure 7.4 (but note that subsequent onset components should have equally affected latency differences for both the EARLY and the LATE condition). Apparently, at this point a more fine-grained data analysis is needed to conclusively quantify the observed effects. Hence, additional analyses relative to potential boundary positions within the stimuli were conducted to meet this need and to be able to draw reliable conclusions regarding the presence of CPS effects in response to the stimuli presented.

7.3.2 Analyses relative to NP offset

Additional analyses were conducted relative to the offset of the first name (or noun phrase, henceforth, NP1 offset), representing the early boundary position, and to the offset of the second name (NP2 offset), representing the late boundary position. Instead of the previous prestimulus baseline ERP epochs were now adjusted to a baseline of 50 ms prior to NP offset. This bears the additional advantages that (1) there is an equal distance between the baseline and the time window used for the statistical analysis at each boundary position under investigation and (2) the relatively short baseline prior to NP offset allows compensating for potential differences in the onset components characterizing the first 400 ms of stimulus processing (see Figure 7.4).

The time window for the statistical analyses relative to NP offset was defined as 100 to 300 ms after NP offset. Given that acoustic cues triggering boundary perception (i.e., final lengthening and pitch change) are already available prior to the offset, a time window starting 100 ms after NP offset should be suitable to evidence the positivity in mean amplitudes signaling IPB processing (see, e.g. Bögels et al., 2010; Pauker et al., 2011). Although CPS effects have been found to peak around 300 to 500 ms after NP offset (e.g., Bögels et al., 2010; Pauker et al., 2011), the time window chosen here ends earlier to avoid an influence of the subsequent P200. Given that pause duration ranges from 250 to 322 ms (see Table 7.2), it is obvious that the P200, a positivity peaking around 200 ms after pause offset, cannot be held responsible for amplitude differences found within the chosen time window.

Separate ANOVAs were conducted for lateral and midline electrode sites including the same topographical levels as in the analysis relative to stimulus onset. Instead of the factor Time window and Condition, the analyses now contained the factors

Position (i.e., either NP1 offset or NP2 offset) and Boundary (boundary status, either with or without IPB) with two levels each. All significant amplitude differences involving the factors Position and/or Boundary are reported and significant interactions with these factors were further analyzed with separate ANOVAs. Where appropriate, a correction according to Greenhouse and Geisser (1959) was applied and reported as the corrected significance.

For lateral sites, an ANOVA including the factors Position, Boundary, Region, and Hemisphere revealed a statistically significant interaction of Position x Boundary x Region x Hemisphere [$F(2,34) = 9.69$, $p < .01$], as well as significant interactions of Position x Boundary x Region [$F(2,34) = 3.83$, $p < .05$], Position x Hemisphere [$F(1,17) = 5.33$, $p < .05$] and Position x Region [$F(2,34) = 7.13$, $p < .01$]. For midline electrodes, an ANOVA including the factors Position, Boundary, and Electrode revealed a significant interaction of Position x Boundary x Electrode [$F(3,51) = 5.41$, $p < .01$] and a significant interaction of Boundary x Electrode [$F(3,51) = 12.87$, $p < .001$].

To test the respective interactions, subsequent statistical analyses were carried out (1) for each position (i.e., NP1 or NP2 offset) and (2) for each boundary status (i.e., with or without IPB) separately. ANOVAs for lateral and midline electrode sites for the early position (NP1 offset) did not reveal significant effects involving the factor Boundary, apart from an interaction of Boundary x Region x Hemisphere for the lateral sites [$F(2,34) = 4.96$, $p < .05$]. Since subsequent two-way ANOVAs for each level of Region and Hemisphere did not reveal any effects for the factor Boundary, this effect was disregarded. Thus, additional statistical analyses suggested no differences between stimuli with and without an IPB at the early position (offset NP1).

In contrast, subsequent ANOVAs at the late position (NP2 offset) revealed main effects of Boundary at lateral sites [$F(1,17) = 6.45$, $p < .05$] and midline electrodes [$F(1,17) = 5.61$, $p < .05$], as well as interactions of Boundary x Region [$F(2,34) = 12.83$, $p < .001$] and Boundary x Electrode [$F(3,51) = 25.68$, $p < .001$], respectively. Hence, a clear difference between stimuli with and without an IPB is present at the late position (NP2 offset).

To determine the topographical position of this effect, subsequent one-way ANOVAs were conducted for each region (lateral sites) and accordingly electrode (midline) testing the aforementioned interactions. Significant main effects of Boundary were revealed for the lateral electrodes at anterior [$F(1,17) = 10.34$, $p < .01$] and central [$F(1,17) = 8.67$, $p < .01$] regions and at the midline electrodes Fz [$F(1,17) = 15.71$, $p < .01$] and FCz [$F(1,17) = 9.86$, $p < .01$], suggesting a fronto-central distribution of the CPS observed for the IPB at the late boundary position.

Regarding the boundary status, ANOVAs for lateral and midline electrode sites for epochs containing no boundary cues [that is, without IPB, either at the early (NP1 offset) or late (NP2 offset) position] did not show significant effects involving the factor Position, apart from an interaction of Position x Region x Hemisphere for the lateral sites [$F(2,34) = 4.41$, $p < .05$]. Since subsequent ANOVAs for each level of Region and Hemisphere did not reveal any effects for the factor Position, this effect was disregarded. Thus, this control comparison suggested no relevant differences between the early position and the late position when no IPB is present.

Crucially, significant differences were obtained comparing epochs with IPB at the early position (NP1 offset) and at the late position (NP2 offset): at lateral electrode sites, a main effect of Position [$F(1,17) = 7.18$, $p < .05$] and an interaction of Position x Region [$F(2,34) = 3.97$, $p < .05$] were present. Analyses for midline electrodes revealed a marginally significant effect of Position [$F(1,17) = 4.37$, $p = .05183$] and an interaction of Position x Electrode [$F(3,51) = 5.46$, $p < .05$]. Hence, the direct comparison between ERPs in response to an IPB at the early position and to an IPB at the late position confirmed the difference between the conditions found in the initial analysis relative to stimulus onset.

To further determine the topography of this difference, subsequent ANOVAs were conducted testing the interactions of Position x Region (lateral sites) and accordingly Position x Electrode (midline). Significant main effects of Position were revealed for the lateral electrodes at anterior [$F(1,17) = 8.15$, $p < .05$] and central [$F(1,17) = 9.57$, $p < .01$] regions and at the midline electrodes Fz [$F(1,17) = 9.30$, $p < .01$] and FCz [$F(1,17) = 6.91$, $p < .05$], supporting the notion of a fronto-central distribution of the CPS effect.

Taken together, the additional statistical analyses confirmed the effect suggested by the initial broader analysis relative to sentence onset (with a prestimulus baseline), but differed from the analysis with a baseline covering the obligatory components of stimulus onset: for the ERPs time-locked to the offset of the critical NPs, statistical differences between the EARLY and the LATE condition were only obtained at the late boundary position (NP2 offset), whereas no differences were present at the early boundary position (NP1 offset). Moreover, epochs with IPB significantly differed as a function of the boundary position, whereas no such amplitude difference could be found for the respective epochs without IPB. Subsequent analyses resolving interactions with the factor Region revealed a fronto-central distribution of the CPS effect observed for the IPB at the late boundary position.

7.4 Discussion

Here, we tested whether the occurrence of the CPS depends on the position of the IPB in the stimuli. In both conditions, the IPB was clearly signaled by three acoustic cues, namely a pitch rise, final lengthening, and a pause. The conditions only differed in regard to the position of the IPB: in the EARLY condition, the IPB already occurred after the first in a list of three names, while in the LATE condition, the boundary occurred after the second name. The results showed that a typical CPS is only elicited in response to a late IPB. When the IPB occurred early in the stimulus material, however, no positive shift was observed. Hence, we found a positional effect which demonstrates that the occurrence of a CPS-like pattern depends on contextual factors such as the position of a prosodic boundary within an utterance.

Given that the occurrence of the CPS indicates prosodic phrase boundary processing we suppose that the prosodic cues, which were unequivocally present in both conditions, were processed in different ways, depending on their position in the utterance. For the LATE condition, the interpretation is straightforward: in line with previous ERP research on prosodic boundary processing (e.g., Steinhauer et al. 1999, see above), the CPS occurs as a marker of IPB processing. The fronto-central scalp distribution matches previous CPS findings. Though the topography of the CPS varies to some extent over studies—presumably depending on the stimuli used (see, e.g., Pannekamp et al., 2005 for different scalp distributions depending on the material used)—CPS effects have been reported not only with a broad distribution (e.g, Kerkhofs et al., 2007; Steinhauer et al., 1999), but also with a fronto-central distribution (e.g., Itzhak et al., 2010). Interestingly, Pannekamp et al. (2005) also found a fronto-central maximum of the CPS when they tested participants with so-called jabberwocky sentences. Since the only content words in the stimuli used here were six proper names, it may well be that the stimuli were processed in a comparable way as stimuli without semantic content, but with appropriate function words and morphemes. Moreover, the latency of the obtained CPS effect is in line with previous studies, where the positive shift has been described to start almost immediately with the end of the preboundary utterance part (i.e., after the onset of the pause; see, e.g., Bögels et al., 2010; Itzhak et al., 2010; Pauker et al., 2011) and to last around 500 to 700 ms (e.g., Pauker et al., 2011, see also above).

Notably, previous ERP studies always used long sentences as stimulus material to investigate prosodic boundary processing. To our knowledge, the current study is the first that demonstrates that the CPS can also be elicited for boundaries in short, non-sentential sequences. Since behavioral studies (e.g., Kentner & Féry, 2013; Lehiste, 1973; Wagner, 2005) have also used this kind of coordinate structure to investigate prosodic phrasing in production and perception, this finding is further evidence for the CPS as an indication of prosodic boundary processing.

In the EARLY condition, two analyses (time-locked to stimulus onset with (1.) a prestimulus baseline and (2.) a baseline covering stimulus initial onset components) came to differential results, see above. As mentioned earlier, the investigation of prosodic boundary processing virtually always comes with critical latency differences in the stimulus material. Therefore we conducted a more sophisticated analysis time-locked to the boundary position (offset of critical NP) in addition to the time-locking to stimulus onset that allows to a) compensate for these inherent latency differences between conditions and b) disentangle the CPS from post-boundary onset components (P200). This analysis confirmed the absence of a CPS in response to the IPB right after the first word (NP1). This is surprising because the boundary cues did not differ in strength from the cues that were present in the LATE condition. As the CPS generally occurs whenever a major prosodic boundary is processed—independent of the segmental, lexico-semantic, or syntactic content

(Pannekamp et al., 2005), we assume that the prosodic cues that were present in the EARLY condition—pitch rise, final lengthening, and pause—were not effectually used for prosodic phrasing and hence did not elicit a CPS as in the LATE condition. In other words, we do not find an effect of on-line boundary processing in the EARLY condition, because the prosodic changes seem not to be interpreted as cues to an IPB. How can this difference in processing be explained? It is assumed here that the crucial difference between the ERP patterns in the EARLY and the LATE condition (and, importantly, also between our EARLY condition and the stimulus material used in previous research on the CPS) lies in the shortness of the first IP. How can we account for an influence of phrase length on the processing of the boundary cues? Below, two possible lines of argumentation will be sketched.

First, the prosodic changes in the EARLY condition may not have been processed as cues to an IPB because there was not enough previous prosodic information available to evaluate them as IPB cues. This reasoning would be in line with behavioral studies demonstrating an influence of the magnitude of a previous prosodic boundary on boundary perception (e.g., Carlson et al., 2001; Clifton et al., 2002; Wagner & Crivellaro, 2010). Remember that these authors argue that prosodic boundary cues are always processed relative to other, previously processed boundaries. In the EARLY condition of our study, no such benchmark is available to the listener when the IPB is encountered, whereas in the LATE condition, in contrast, a weaker prosodic boundary (signaled by the moderate pitch rise at the end of the first name) has already been processed once the IPB occurs. Moreover, as prosodic information is not only available at boundary positions but unfolds over time, the previously processed prosodic context in general may serve as a reference system for boundary perception. Our results would hence in addition reflect a length effect: the processing of local boundary cues relies on previously processed prosodic information, which in our case unfolds during the perception of the longer constituents in the LATE condition. In the EARLY condition, no such or not enough previous contextual information (e.g., information on segmental duration and pitch variation) is available to interpret the prosodic boundary cues as such during on-line processing. This length effect seems at first glance inconsistent with the behavioral results of Clifton et al. (2006). Remember that the authors found a larger effect of boundary perception for stimuli with short constituents as compared to stimuli with long constituents and argued that boundaries after short constituents are more informative to the listener. In contrast, we found an effect of boundary processing only at the late boundary position, when a longer constituent precedes the boundary. However, this contrast may be easily explained by the differences in the experimental design. Clifton et al. (2006) ascribe their finding to the fact that after long constituents a prosodic break may be inserted for reasons irrelevant to syntactic parsing (i.e., speech fluency). As the coordinate structures we used were in general rather short—even in the LATE boundary condition the first constituent consisted of no more than five syllables—this reasoning does not necessarily hold for our material and thus it may not be appropriate to expect the type of length effect Clifton et al. (2006) describe. Hence, despite the reversed direction of the length effect, our results are in accordance with the findings of Clifton and colleagues and could either mirror a mere length effect or be interpreted in line with the findings on the impact of relative prosodic boundary strength (see above). Accordingly, our results are consistent with a non-local account for prosodic boundary processing assuming a context-dependent interpretation of prosodic boundary cues.

Second, the missing CPS in the EARLY condition may be due to an unnecessity for chunking at the early boundary position. Prosodic boundaries enable the listener to chunk the incoming auditory signal into larger units and may hence help to reduce processing costs and to guide the parser. Remember that in the EARLY condition, the boundary cues are encountered when listeners have only perceived a minor part of the utterance (in fact, only two syllables, i.e., the first proper name). Therefore, there may simply be no need to chunk this word into a larger (prosodic) unit—a cognitive process that may be reflected by the CPS. This would imply that the significance of the CPS goes beyond the pure detection or encounter of prosodic boundary cues. In fact, our data support the idea that the CPS is not only mirroring perceptual processes. Rather, it has a linguistic or cognitive relevance signaling the use of prosodic boundary information during on-line processing. Note that this reasoning holds for the notion of unnecessary chunking as well as for the suggested account of lacking prosodic context.

Crucially, one has to keep in mind that the initial analysis time-locked to sentence onset with a baseline controlling for differences in onset components indicate the presence of a more positive going ERP at the point of the IPB also in the EARLY condition. This highlights the susceptibility of ERP analyses to the choice of baseline and time window parameters especially in speech processing and the necessity of a proper stimulus design that allows for a time-locking to the critical events in the speech stream.

Future research is necessary to clarify under which conditions a CPS may be elicited even at the first boundary position or after short IPs. For example, one could gradually enlarge the first noun phrase by using polysyllabic names or by adding a determiner or a modifying adjunct. However, at least in the latter case one has to keep in mind that adding more material to the IP may lead to an additional (weaker) prosodic boundary. With the material used in our study we clearly cannot disentangle if the contextual information necessary to elicit a CPS is purely prosodic in nature or whether also additional syntactic or lexico-semantic information may add to the visibility of a CPS in the EARLY condition. Given the findings of Pannekamp et al. (2005) we speculate that linguistic domains other than prosody have a minor influence on the CPS. However, due to the absence of a specific task in our experimental setting, it may well be that listeners did not entirely process the syntactic structure of the material. A task forcing participants to resolve the syntactic ambiguity may hence lead to enhanced CPS effects. Potentially, this could imply the occurrence of a CPS in the EARLY condition.

7.5 Conclusions

The present study yields two pieces of evidence for the incremental processing of prosodic information. First, the immediate integration of prosodic boundary cues is reflected by the CPS elicited in the LATE condition. Second, contextual prosodic information may also have an immediate influence on the processing: in the EARLY condition, the use of prosodic boundary cues seems not be warranted by the preboundary context, either due to missing benchmark prosodic information or because it is not necessary from a cognitive resource viewpoint. Therefore, no CPS-like ERP pattern occurs. This shows, in turn, that the occurrence of a CPS does not reflect the brain's response to acoustic changes which may indicate an IPB, but rather that it mirrors the integration of available prosodic boundary information into the parsing process, that is, it signals the use of prosodic boundary cues for sentence comprehension. In conclusion, we have shown that a CPS is not necessarily elicited whenever the relevant prosodic boundary cues are present. Instead, the occurrence of the CPS was influenced by the IPB position, which was correlated with differences in the length of the preceding constituent and in the occurrence of an earlier boundary in the stimulus material. Further research is needed to determine the exact nature of the apparent impact on CPS occurrence. The result can be interpreted in line with a non-local account of boundary processing, because previously processed information has an immediate impact on the processing mechanism. In addition, by using electrophysiology we find evidence for an immediate integration of prosodic cues into the parsing of an utterance as long as this is affirmed by the previously processed context. Regarding the functional relevance of the CPS, this study yields further evidence that the CPS does not reflect pure signal detection, but rather mirrors the use and integration of prosodic boundary information during on-line spoken language comprehension.

Acknowledgements

This research was funded by a grant from the German Science Foundation, priority program 1234, to Isabell Wartenburger and Barbara Höhle (FR 2865/2-1; HO 1960/13-1). Isabell Wartenburger is supported by the Stifterverband für die Deutsche Wissenschaft (Claussen-Simon-Stiftung). We thank Ruben van de Vijver for helpful comments on earlier versions of this manuscript.

8 How pitch change and final lengthening cue boundary perception in German: converging evidence from ERPs and prosodic judgments[2]

Abstract

This study examines the role of pitch and final lengthening in German intonational phrase boundary (IPB) perception. Since a prosody-related event-related potential (ERP) component termed *Closure Positive Shift* reflects the processing of major prosodic boundaries, we combined ERP and behavioral measures (i.e., a prosodic judgment task) to systematically test the impact of sole and combined cue occurrences on IPB perception. In two experiments we investigated whether adult listeners perceived an IPB in acoustically manipulated speech material that contained none, one, or two of the prosodic boundary cues.

Both ERP and behavioral results suggest that pitch and final lengthening cues have to occur in combination to trigger IPB perception. Hence, the combination of behavioral and electrophysiological measures provides a comprehensive insight into prosodic boundary cue perception in German and leads to an argument in favor of interrelated cues from the frequency (i.e., pitch change) and the time (i.e., final lengthening) domain.

[2] This chapter has been published as:
Holzgrefe-Lang, J., Wellmann, C., Petrone, C., Räling, R., Truckenbrodt, H., Höhle, B., & Wartenburger, I. (2016). How pitch change and final lengthening cue boundary perception in German: converging evidence from ERPs and prosodic judgements. *Language, Cognition and Neuroscience, 31*(7), 904-920.

8.1 Introduction

Spoken language comprehension benefits from the prosodic information provided in the speech signal, chunking utterances into prosodic phrases of different strengths (see, e.g., Nespor & Vogel, 1986; Selkirk, 1986; Truckenbrodt, 2007a). Moreover, prosodic phrasing is also central to language acquisition, as already young infants rely on the prosodic structure of their native language to segment the continuous speech signal and to extract meaningful units (see, amongst others, Hirsh-Pasek et al., 1987; Nazzi et al., 2000; Soderstrom et al., 2003). Infants as well as adults benefit from a close syntax-prosody mapping; more precisely, the edges of major prosodic phrases, so-called Intonational Phrase boundaries (IPBs), usually coincide with syntactic boundaries (e.g., Downing, 1970; Selkirk, 2005; for German: Truckenbrodt, 2005). Accordingly, numerous behavioral studies on adult spoken language comprehension demonstrated an influence of prosodic boundary processing on syntactic analysis (see, for a review, Cutler et al., 1997 and, amongst others, Kjelgaard & Speer, 1999; Price et al., 1991; Sanderman & Collier, 1997; Schafer, 1997). Along with (optional) pauses, an IPB—as the largest domain in the prosodic hierarchy (see, e.g., Selkirk, 2005; Truckenbrodt, 2005)—is mainly signaled by two further prosodic cues: across different languages, pitch change (i.e., a pitch rise or pitch fall indicating the presence of a boundary tone and/or a pitch reset) and final lengthening (i.e., an increase in the duration of the segments immediately preceding the boundary) were identified as main prosodic boundary cues (see, amongst others, Hirst & Di Cristo, 1998; Vaissière, 1983; Wightman, Shattuck-Hufnagel, Ostendorf, & Price, 1992; and also Wagner & Watson, 2010 for a review; for German: Gollrad et al., 2010; Kentner & Féry, 2013; Kohler, 1983; Peters et al., 2005). While the use of the prosodic cues is a universal phenomenon, the cue weighting, that is, the relative importance of one cue (i.e., final lengthening) as compared to another one (i.e., pitch change), may depend on exposure to the target language and thus differ across languages. Recent language acquisition research focused on the development of a language-specific weighting of the key prosodic cues (see, e.g., Johnson & Seidl, 2008, Seidl, 2007, Seidl & Cristià, 2008, Wellmann et al., 2012). With regard to adult processing, research on the role of a specific boundary cue or cue combination in sentence comprehension is sparse. Even though inferences on cross-linguistic variation in cue weighting are restricted by the limited number of studies systematically investigating the impact and relative importance of specific cue subsets, we will, in what follows, first review the relevant behavioral studies from different languages. Since the present study is, to the best of our knowledge, the first to address this topic using a behavioral task in combination with an electrophysiological measure, the remaining part of the introduction will be dedicated to the presentation of event-related potential (ERP) research on IPB perception and possible implications regarding prosodic cue processing.

In a pioneering study on prosodic boundary cue perception in English, Streeter (1978) used ambiguous algebraic expressions (e.g., "(A plus E) times O" vs. "A plus (E times O)", example taken from Streeter, 1978) that were manipulated by interchanging the pitch contour and the duration pattern (i.e., final lengthening) from the two alternative phrasings. Participants were asked to decide on the boundary location (i.e., which of the above cited bracketings was conveyed by the speaker). Streeter found that both final lengthening and pitch change were employed by the listener as single cues, since the tendency to accept the alternative phrasing was higher in both single cue conditions (final lengthening: about 46 %; pitch change: about 39 %) as compared to the original stimulus (about 24 %). However, only the combination of both cues completely shifted the listeners' decision from the original phrasing to the alternative phrasing (with an acceptance rate of about 72%). Hence, Streeter (1978) assumes an additive effect of the two prosodic cues.

Scott (1982) focused on the two temporal boundary cues, that is, final lengthening and pause. She used short English sentences that contained a string of three proper nouns separated by different conjunctions (e.g., "Kate and Pat or Tony will come"). Two alternative groupings of the noun phrases were possible, resulting either in a major prosodic boundary after the first or after the second noun. The stimuli were edited in a stepwise manner by systematically increasing or decreasing the amount of final lengthening and pause duration. The combination of final lengthening and pause duration was found to be an effective boundary cue, increasingly shifting the listeners' decision towards the alternative phrasing. However, listeners also identified a prosodic boundary if only an extended pause was present. The impact of final lengthening as a sole cue, on the contrary, was not tested in the study.

Aasland and Baum (2003) also investigated the role of final lengthening and pause in complex conjunct phrases (namely, “(pink and black) and green” vs. “pink and (black and green)”) using a 5-step continuum of natural speech manipulations. Their results confirmed that final lengthening and pause are most effective in combination. However, a sufficiently large final lengthening cue also shifted the participants’ decision towards the alternative phrasing, whereas—in contrast to the findings by Scott (1982)—even the longest pause cue was insufficient in the absence of final lengthening. The step size for the pause manipulation was 40 ms (i.e., the longest pause was only 160 ms long), although pauses up to 256 ms were found in the natural speech stimuli. Considering that the pause duration continuum for the manipulated stimuli in Scott (1982) ranged from 0 up to 562 ms, this may explain the inconsistent findings.

Recently, Zhang (2012) compared the perception of pitch change, final lengthening, and pause cues in English and Mandarin Chinese. For each language, an ambiguous utterance pair that—depending on the presence of a prosodic boundary—contained lists of either two or three nouns (e.g., “turkey, salad, and coffee” vs. “turkey-salad and coffee”) was acoustically manipulated by altering the rime duration of the syllable preceding the potential boundary (here: “-key”) and pause duration, respectively, in a 5-step continuum. The pitch cue was not manipulated in a stepwise fashion, but represented the pitch contour from the naturally produced utterances either with or without prosodic boundary. Zhang (2012) found that English and Mandarin Chinese listeners made use of all three investigated cues to perceive prosodic boundaries in their native language, with final lengthening being the weakest cue in both languages. The remaining two cues were weighted differently: English listeners relied more on pause than on pitch contour change, whereas Chinese listeners weighted pitch (i.e., pitch reset) more heavily than pause. The difference in the reliance on pitch information is supposed to be associated with the phonemic status of pitch in the Chinese’s phonological system (as compared to non-tonal languages like English).

In another recent study on Mandarin Chinese (Yang et al., 2014), boundary cues were systematically removed from sentences that originally contained two intonational phrases (and hence an IPB in between). As expected, the proportion of boundary detection was lowest when all three boundary cues were removed. Stimuli in which either final lengthening or the pitch reset was preserved (whereas the other two respective cues were removed) yielded comparably higher detection rates. However, detection rates were significantly lower than for stimuli with a preserved pause only, as well as for stimuli containing pause and final lengthening information (whereas the pitch cue was removed). Results for the latter two conditions did not differ from the natural control condition containing all boundary cues. Thus, unlike Zhang (2012), Yang et al. (2014) claim that pause is the most powerful IPB cue in Mandarin Chinese, whereas pitch and final lengthening are weaker and, in addition, perceptually equivalent cues. The pause duration employed by Yang et al. (2014) averaged out at 270 ms and represented the duration observed in the natural speech stimuli. Zhang (2012), using ambiguous lists instead of whole sentences, manipulated her stimuli gradually and inserted pauses that were at maximum 80 ms long, although pauses of more than 300 ms occurred in the production of the natural stimuli with IPB. Hence, Yang et al. (2014) assume that Zhang's (2012) composition of stimuli reduced the impact of the pause cue and thus ascribe the diverging results to this difference in pause duration.

Regarding the perception of boundary cues in German, a study by Gollrad et al. (2010) shed light on the role of the pitch contour as compared to durational cues (i.e., the combined impact of final lengthening and pause) in the perception of locally ambiguous sentences. In a sentence completion task, listeners were presented with sentence fragments in two prosodic conditions, either with neutralized durational distinctions (i.e., without pause and final lengthening cues) or with a flattened pitch contour. When the pitch contour was flattened, participants still selected the original sentence continuation successfully, while they failed to do so when the pitch contour was present but the durational cues were neutralized. This suggests that durational cues are perceptually more relevant for German IPB processing than pitch. However, the study design does not allow for distinguishing between the impact of pause and final lengthening cues, since they were manipulated simultaneously.

To summarize, behavioral studies indicate that pause is a very salient cue, and seems to be the most dominant boundary cue, at least in non-tonal languages.

Although there is consistent evidence that both final lengthening and pitch change also contribute to the perception of an IPB, their specific role and effectiveness as single cues remains unclear. With regard to the production of prosodic boundary cues in German spontaneous speech, a corpus study (Peters et al., 2005) found that although cue combinations are frequent, 31.5% of all turn-internal boundaries were only marked by one of the three phrase boundary cues (20.8% pitch change, 9.4% final lengthening, and 1.3% pause). Moreover, the combination of pitch change and final lengthening occurred predominantly among the possible combinations of two boundary cues: While 24.6% of all boundaries were marked the combination of pitch change and final lengthening, only 8.4% were marked by the co-occurrence of final lengthening and pause and only 4.9% by pitch change and pause. Therefore, the present study is concerned with the contribution of pitch change and final lengthening to German phrase boundary perception, both as single cues and in combination. Examining pitch and final lengthening cues only bears the advantage that the two rather subtle cues can be investigated irrespective of the impact of the pause cue, especially since Peters (2005) found for German that the presence of a pause can mask effects of other boundary cues. Our approach is thus orthogonal to Scott (1982) and Aasland and Baum (2003) who specifically compared the two durational cues or to Gollrad et al. (2010) who did not subdivide their notion of duration as a boundary characteristic into specific prosodic cues.

The behavioral studies presented so far are inherently limited in that off-line methods such as judgments and reaction time data only display the result of a complex perceptual operation once the global prosodic structure (i.e., the whole stimulus) has been processed. On-line methods with a high temporal resolution, like the event-related potential (ERP) technique, capture the immediate impact of a local prosodic cue at the potential boundary location. A further contribution of the present study is hence the combined use of complementary on-line and off-line research techniques, namely a prosodic judgment task and ERP recordings.

ERPs have been found to be a useful tool in addressing questions on the integration of prosodic information in auditory processing (see, e.g., Eckstein & Friederici, 2006) and, more specifically, on the time course of prosodic phrase boundary processing (see, for a review, Bögels et al., 2011). In a study on German sentence comprehension, Steinhauer et al. (1999) presented stimuli that contained either one or two IPBs. In response to each IPB they found a broadly distributed ERP component with a positive going waveform that was largest at central and parietal electrode sites. Due to the co-occurrence with the closure of a major prosodic phrase, it was termed Closure Positive Shift (CPS). To date, the CPS has been established as an indicator of IPB processing in various languages: in German (e.g., Holzgrefe et al., 2013; Scott, 1982; Pannekamp et al., 2005) English (e.g., Itzhak et al., 2010; Pauker et al., 2011; Peter et al., 2014), Dutch (e.g., Bögels et al., 2010; Kerkhofs et al., 2007), Swedish (Roll & Horne, 2011), Japanese (Wolff et al., 2008), Mandarin Chinese (Li & Yang, 2009), and Korean (Hwang & Steinhauer, 2011). In particular, several studies contributed to unravel the underlying mechanisms that trigger the CPS: In a series of experiments, Pannekamp et al. (2005) provided evidence for a truly prosodic origin of the CPS response by demonstrating that the CPS can be elicited in the absence of semantic, syntactic, or segmental information, albeit with varying scalp distributions. Moreover, Steinhauer et al. (1999) as well as Männel and Friederici (2009) demonstrated that a pause between two intonational phrases is not necessary to elicit a CPS in adults. This clearly indicates that the CPS is not a variant of early auditory evoked potentials that signal the detection of new auditory input (e.g., after a break) and does thus not reflect lower level acoustic processing. In addition, this can be seen as a first hint that IPB perception, as reflected by the CPS, does not depend on the presence of a pause. Finally, the facts that a CPS has even been elicited in the absence of acoustic input (implicit prosody in silent reading, e.g., Hwang & Steinhauer, 2011) and that, under certain conditions, the CPS can be absent even though the acoustic boundary cues are processed (Holzgrefe et al., 2013) strongly point to underlying linguistic processes (i.e., boundary perception) as compared to purely acoustic processing (i.e., acoustic signal detection).

The present study makes use of the CPS as an indicator of IPB perception and combines electrophysiological and behavioral measures to determine the impact of specific prosodic cues on boundary perception in German. Importantly, it is one of the first ERP studies on the impact of pitch change and final lengthening using systematically manipulated stimulus material. With regard to the CPS, the reasoning goes as follows: if pitch change and final lengthening are sufficient to trigger boundary perception as sole cues or in combination, a CPS should be elicited for the respective experimental conditions as compared to a baseline condition containing no prosodic boundary cues. If, in contrast, pitch change and/or final lengthening are not effective as single cues, there should be no qualitative difference in the brain response to the respective conditions. Based on the previous behavioral and ERP results we hypothesize that a CPS should occur in response to combined pitch and final lengthening cues. Concerning the impact of single prosodic cues empirical evidence so far is inconsistent and, especially with regard to ERPs, very sparse. Crucially, if a CPS is elicited for more than one condition, CPS amplitude differences may shed light on additive functioning and the relative importance of the prosodic cues. Though we expect both judgments and ERPs to reflect IPB perception concordantly, complementary results may also provide an additional insight into prosodic boundary cue processing.

8.2 Methods

8.2.1 Participants

Twenty-eight participants (22 women, age range 20–30 years, mean age 23.5 years) took part in Experiment 1 and 30 participants (20 women, age range 18–29 years, mean age 21.7 years) in Experiment 2. All participants were native speakers of German with no reported hearing or neurological disorders and were right-handed, as attested by a German version of the Edinburgh Handedness Inventory (Oldfield, 1971). They were students of the University of Potsdam and received course credits or reimbursement for their participation. All participants took part in the experiments after giving informed consent and the study was approved by the local ethics committee.

In Experiment 1, a further two participants were excluded from the final data analysis because of excessive artifacts in the electroencephalogram (EEG). Participants were excluded if the EEG data did not meet the criterion of at least 70% artifact-free trials per condition. In Experiment 2, the data of one additional participant were excluded from the final data analysis, because the participant's behavioral response exhibited a strong bias in that only one of the two buttons involved in the forced-choice task was pressed in over 99% of the trials.

8.2.2 Material

Consistent with previous studies (e.g., Aasland & Baum, 2003; Scott, 1982) the stimulus material contained strings of three proper nouns separated by conjunctions. The stimuli were recorded in an anechoic booth equipped with an AT4033a audio-technical studio microphone, using a C-Media Wave soundcard at a sampling rate of 22050 Hz with 16 bit resolution. A young female German native speaker from the Brandenburg area was provided with a written list of the stimuli in which two grouping alternatives were indicated by bracketing as in (8). Each stimulus recording was preceded by the same context question (*Wer kommt?* 'Who is coming?') read by the experimenter. The speaker was instructed to read the name triples in such a way that the experimenter and an independent second listener, both being naïve to the given bracketing, were able to conceive the bracketing and add it to a plain version of the stimulus list. This ensured that the stimuli were produced with the intended prosodic structure.

(8) a. (Moni und Lilli und Manu)

b. (Moni und Lilli) (und Manu)

The sequences consisted of three disyllabic, trochaic German proper nouns that were coordinated by *und* ('and'). Each proper noun is a syntactic noun phrase and gives hence rise to a phonological phrase (PhP), set off by a PhP boundary from the other names (Gussenhoven, 1992; Truckenbrodt, 1999; 2007a). Both sequences contain the same string and are ambiguous in that all three names can be grouped together as shown in (8a) or, for example, the first two names are grouped together and the final one is apart as shown in (8b). The disambiguation of the alternative groupings employs the next higher level of the prosodic hierarchy, that is, the intonational phrase (Selkirk, 2005; Truckenbrodt, 2005). Thus, sequences of type (8a) are produced as a single intonational phrase and hence do not contain a sequence-internal IPB, whereas sequences of type (8b) are produced with an IPB after the second name as illustrated in (9).

(9) a. Without IPB: $[(\text{Moni})_{PhP}$ $(\text{und Lilli})_{PhP}$ $(\text{und Manu})_{PhP}]_{IP}$

b. With IPB: $[(\text{Moni})_{PhP}$ (und Lilli)PhP$]_{IP}$ # $[(\text{und Manu})_{PhP}]_{IP}$

Six different name sequences were constructed: Six German proper nouns (*Leni, Lilli, Manni, Mimmi, Moni,* and *Nelli*) occurred once in the first and once in the second position, three additional names (*Lola, Manu,* and *Nina*) occurred each twice in the utterance-final position (see Appendix A.1 for a complete list of experimental stimuli). All proper nouns were composed of sonorant sounds to facilitate a reliable measurement of the fundamental frequency (f0)—the acoustic correlate of the pitch contour.

A total of 60 recordings, namely five tokens of each name sequence in both prosodic conditions (without and with IPB, see Example 9a and 9b), entered the acoustic analysis. Acoustic analyses using PRAAT software (Boersma & Weenink, 2010) confirmed that sequences with IPB clearly manifest the acoustic correlates of the three main prosodic boundary cues: a pitch rise leading to a high boundary tone occurred on the second name and was followed by a silent pause. The mean pause duration was 532 ms, compared to no pause at this position in sequences without IPB. To assess the pitch change, the second name was decomposed into four intervals corresponding to the phonetic segments, that is, the single consonantal and vocalic parts of the signal. F0 was measured at the position of minimum pitch (on the first segment) and maximum pitch (i.e., a high boundary tone present on the final vowel). The difference between the f0 minimum on the first segment and the f0 maximum on the final vowel was used to calculate the pitch rise. On average, this pitch rise was 2.3 times greater in sequences with IPB compared to sequences without IPB. Final lengthening occurred on the pre-boundary syllable and was assessed by measuring the length of the final vowel of the second name in both prosodic conditions. Mean vowel duration was about 1.8 times longer in sequences with IPB compared to sequences without IPB (see Table 8.1).

The parameters observed in the acoustic analysis served as a reference for the subsequent acoustic manipulations. In particular, mean values pertaining to pitch change and final lengthening (as employed by this speaker) were adopted in generating new experimental conditions. Since previous work (Holzgrefe et al., 2013) had already shown that coordinated structures with IPB after the second name are suitable to assess boundary processing on the neurophysiological level, stimuli with a natural IPB, that is, sequences containing all three boundary cues were not included as experimental stimuli.

Table 8.1 | Mean values (range) of the acoustic correlates of prosodic boundary cues (measured on and following the second name) in the 60 original stimuli.

Prosodic boundary cue	Acoustic correlate	Without IPB	With IPB
Pitch change (rise and high boundary tone)	f0 rise in Hz	97 (69–130)	223 (183–266)
	Maximum f0 in Hz	282 (252–309)	401 (357–431)
Final lengthening of the	Final vowel duration in ms	96 (74–127)	168 (144–200)
Pause	Pause duration in ms	0	532 (363–768)

Acoustic cue manipulation and experimental conditions

The purpose of the present study is to determine the specific impact of pitch change and final lengthening separately and in combination. To test which boundary cue or cue combination is necessary and/or sufficient to trigger the perception of an IPB, the stimuli must differ from a baseline condition without boundary cues (i.e., not triggering a CPS /IPB perception) only with respect to the critical boundary cue or cue combination. Moreover, local cue manipulations were a necessary prerequisite for analyzing the ERPs, because ideally the stimuli should here only differ from the baseline condition within a predefined critical region. Therefore, we decided to approach acoustic manipulation from the side of stimuli with a flat structure: we started with name sequences without an IPB and locally added the crucial boundary information to avoid a potential influence of additional cues that may contribute to IPB perception. Hence, experimental effects can be clearly allocated to the acoustic properties under investigation. In addition, the chosen acoustic manipulations allowed us to investigate pitch and final lengthening in the absence of a (presumably more salient) pause cue. Further, adding the respective prosodic cues allowed for a thorough control of the phonetic and phonological characteristics of the material and the naturalness of the material was maintained as much as possible (by avoiding the compulsory removal of potential additional acoustic cues).

Therefore, all subsequently described stimuli were systematically constructed from the sequences without IPB (i.e., the future condition NO, see below) by using PRAAT software (Boersma & Weenink, 2010). The experimental stimuli were created by inserting either a pitch cue (conditions PI and PI2) or a final lengthening cue (DUR) or the combination of both (PIDUR and PI2DUR). Following acoustic manipulation, all sequences were resynthesized using the PSOLA function in PRAAT. The acoustic manipulation steps are detailed below.

The *condition NO* served as a baseline condition in both experiments since no IPB cues were present on the second name (see Figure 8.1A). It contained six original sequences without IPB, namely one recording of each name combination, that were scaled to a mean intensity of 70 dB. However, to avoid comparing natural with acoustically manipulated material, the pitch contour of the stimuli was stylized (2 semitones). This transformation reduces the number of pitch points and formed the basis for further acoustic manipulations in all other conditions. The *condition PI* consisted of the stylized stimuli from condition NO with an inserted pitch rise on the second name as a sole boundary cue. The reference values of the fundamental frequency were measured on the second name in the 60 original sequences with internal IPB, namely at the midpoints of the four segments and at the position of the maximum pitch present on the final vowel (see above and Table 8.1). For the manipulation of the pitch contour, pitch points with the mean values at these time points were inserted at the same positions into the stylized sequences of condition NO. The six new stimuli contained a natural sounding pitch rise of 206 Hz (13.16 semitones). Exemplar spectrogram and pitch contour are depicted in Figure 8.1B. The *condition DUR* served to test the impact of final lengthening as a sole cue. Therefore, in the six stylized sequences without IPB, the final vowel of the second name was lengthened to 180%. This factor was chosen because in the natural stimuli, the crucial vowel was on average 1.8 times longer in sequences with IPB than in sequences without IPB (see Table 8.1). Figure 8.1C displays an experimental stimulus from condition DUR. The *condition PIDUR* tested the impact of pitch and final lengthening cues in combination. To this end, a pitch rise was implemented just as described for condition PI and the final vowel of the second name was lengthened to 180% as described for condition DUR. A stimulus containing the combined manipulation of pitch and lengthening is depicted in Figure 8.1D.

The conditions described so far employed local prosodic boundary cues that occur in close vicinity to the IPB position. In other words, acoustic manipulations were only carried out on the two syllables (i.e., the second name) preceding the possible boundary position, while the rest of the stimulus remained identical across conditions. Yet, the examination of the natural stimuli had revealed that a slight pitch rise occurred on the first name in the natural sequences without IPB. This pitch change on the first name was not present in the natural stimuli of the contrasting sequences with IPB. Thus, we assume that it may serve as a prosodic cue in that the presence of a minor pitch rise on the first element of a list could hint at the absence of a larger boundary after the second element of a coordinated list (see also Kentner & Féry, 2013 for a similar finding). To avoid potential inference of this distant opposed pitch cue, two additional experimental conditions were constructed that did not contain the early pitch rise. In particular, the *condition PI2* contained stimuli that exhibited the pitch rise on the second name (see condition PI) and in addition a flattened pitch contour on the first name. Flattening of the pitch contour was carried out similarly to the pitch manipulation on the second name: For both names, reference f0 values were measured at the midpoints of the four segments in stimuli with IPB and subsequently, pitch points with the mean values at these time points were inserted at the same positions into the sequences without an IPB. An exemplar stimulus of the condition is depicted in Figure 8.1F. Accordingly, the six stimuli of *condition PI2DUR* had the same acoustic properties as the stimuli in condition PIDUR except that the pitch contour on the first name was flattened (see Figure 8.1E). In contrast to the local manipulations that were carried out in close proximity to the possible IPB position, the latter two conditions thus contained rather global pitch cues, affecting both the first and the second name of the sequences.

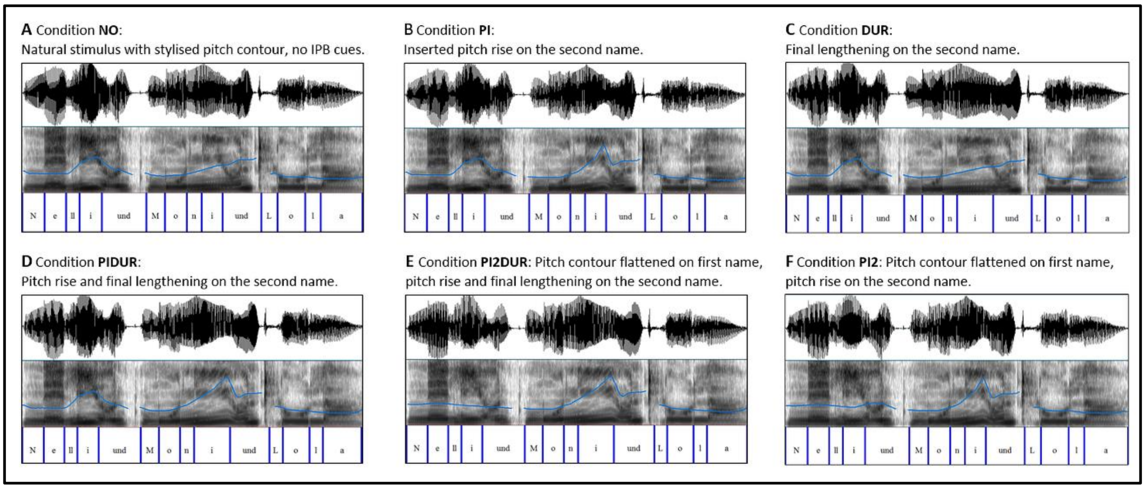

Figure 8.1 | Exemplar waveforms, spectrograms, and pitch contours of the experimental stimuli.

During the experiments, each of the six stimuli was presented 10 times, yielding a total of 60 experimental items per condition. Each experiment consisted of four conditions, with condition NO being present as a baseline in both Experiment 1 and 2. Besides, the condition PI2DUR was employed in both experiments because it represents an additional “upper bound” control condition containing the largest amount of boundary cues across all experimental conditions. Experiment 1 further contained the conditions PI and PIDUR, and hence tested the impact of local pitch change as a sole cue (PI) and in combination with final lengthening (locally: PIDUR, globally: PI2DUR) against the baseline condition NO without boundary cues. The direct comparison of PIDUR versus PI2DUR may allow conclusions to be drawn about the impact of the non-local versus the local pitch cue. Experiment 2, in turn, additionally contained the conditions DUR and PI2. Hence, in Experiment 2 we investigated the impact of final lengthening as a sole cue (DUR), the influence of the global pitch cue as a sole cue (PI2) and the combined occurrence of these two single cues (PI2DUR) against the baseline condition NO without boundary cues.

8.2.3 Procedure

The experiment was performed using Presentation® software (version 14.1; Neurobehavioral Systems, www.neurobs.com). The 240 experimental items were presented aurally (using E-A-RTONE 3A Insert Earphones, Aearo Technologies Auditory Systems, Indianapolis, USA) in a pseudo-randomized order: at most two consecutive trials belonged to the same condition and the same sequence of names never occurred in consecutive trials.

The prosodic judgment task was carried out in a sound-attenuating chamber while the EEG was recorded. Each participant was seated in a comfortable chair and was fitted with an elastic cap (EASYCAP) that held the EEG electrodes. Once the electrodes were fitted, participants were instructed to avoid eye blinking and other body movements during stimulus presentation. In order to minimize eye movements, each trial started with the presentation of a fixation cross in the center of the monitor, lasting for 500 ms. Then, the auditory stimulus was presented. The fixation cross remained visible on the screen for another 700 ms after stimulus offset. Subsequently, the two alternative bracketings (i.e., [X and Y] [and Z] vs. [X and Y and Z]) were depicted on the left and right side of the screen (presentation side balanced across participants) and participants had to decide via button press which of the bracketings better matched the stimulus. Responses were recorded within a 3000 ms time window. To encourage intuitive judgments, however, participants were instructed to respond as soon as the bracketings appeared on the screen. The button press was followed by an inter-stimulus-interval of 3000 ms, before the next trial started with the presentation of the fixation cross. The experimental session included a short practice phase of four trials (one per condition) to ensure that participants were familiar with the judgment task and a short break after half of the experimental trials. The experiment lasted about 30 minutes.

EEG recording and data preprocessing

The EEG was continuously recorded from 29 active Ag/AgCl electrodes (actiCAP, Brain Products, Germany) with a sampling rate of 1000 Hz. The electrodes were placed according to the international 10–10 system (American Clinical Neurophysiology Society, 2006) at the following positions: Fpz, AFz, Fz, F3/4, F5/6, F7/8, FCz, FC3/4, Cz, C3/4, T7/8, CPz, CP5/6, Pz, P3/4, P7/8, POz, PO3/4, Oz. The electrooculogram was recorded from electrodes placed above and below the right eye. Electrical activity was recorded from both the mastoids, and the EEG recording was referenced on-line to the left mastoid and re-referenced off-line to linked mastoid electrodes. Impedances were kept below 5 kΩ.

The EEG data were analyzed using Brain Vision Analyzer (version 2.01; Brain Products, Gilching, Germany). A digital band pass filter ranging from 0.2 to 70 Hz was applied to remove very slow drifts and muscle artifacts, and we also applied a 50 Hz notch filter. Epochs of 2500 ms, starting 200 ms prior to stimulus onset, were extracted from the continuous EEG signal. Eye blinks and eye movements in the epochs were corrected by a computer algorithm (Gratton et al., 1983). All other artifacts were detected semi-automatically, and contaminated segments were excluded from further analysis.

In Experiment 1, the mean number of averaged trials per participant was 53.25 for condition NO (*SD* = 4.37; 88.75%), 52.71 for condition PI (*SD* = 3.92; 87.86%), 52.82 for condition PIDUR (*SD* = 4.73; 88.04%), and 52.79 for condition PI2DUR (*SD* = 4.65; 87.98%). In Experiment 2, the mean number of averaged trials per participant was 56.23 for condition NO condition (*SD* = 2.76; 93.72%), 56.7 for condition DUR (*SD* = 3.02; 94.5%), 57.07 for condition PI2 (*SD* = 2.82; 95.11%), and 56.37 for condition PI2DUR (*SD* = 3.12; 93.94%).

8.2.4 Statistical analysis

Behavioral data

Following Jaeger (2008), the statistical analysis of the categorical judgment data employed logit models with mixed effects (or simply "mixed logit"). Mixed logit models are based on binomial distributions (z-scores) and therefore allow modeling binary decisions like two-alternative forced-choices. Unlike standard logistic regression, the inclusion of mixed effects bears the additional advantage that they cover both fixed and random effects within one analysis. We used the statistical software R (R Core Team, 2013) with the supplied LME4 package (Bates, D. M., Maechler, M., & Dai, B., 2009) which fits mixed logit models by Laplace approximation.

For each experiment, a separate model was run with Condition as fixed effect containing four levels (Exp. 1: NO, PI, PIDUR, and PI2DUR; Exp. 2: NO, PI2, DUR, and PI2DUR). Participants and items were entered as random intercepts. Model fitting always started with the most complex model, that is, with the full factorial set of random effects (random slope adjustments for Condition for participants, items, and repetitions). Backward elimination of non-significant random terms was applied in a stepwise manner; the complex model was reduced by model comparisons using log-likelihood tests (e.g., Baayen, 2008; Baayen, Davidson, & Bates, 2008). Estimates (β), standard errors (*SE*), z-values (z) and the level of significance (*p*) of the final models are reported.

We also verified (1) whether all levels of Condition differed from each other and (2) whether the response score for each level of Condition is different or equal to the chance level. Both analyses were achieved by resetting the intercept of the models: the same model was run different times for each experiment, each time changing the intercept of the model, that is, the level of Condition which is the reference one for pairwise comparisons with all other levels of that factor. This also allowed comparing each intercept (i.e., each level of Condition) to the zero value of the logit function, which represents a probability of 50%. If the intercept is significantly different from zero, it means that its probability is significantly below or above the chance level. An intercept that does not differ from zero in the logit model corresponds to a response score at the chance level. Given that multiple comparisons were made between each level of Condition and the chance level, *p*-values were adjusted according to Holm (1979).

ERP data

To compensate for latency differences in the experimental material (due to final lengthening occurring in two out of the four experimental conditions), the statistical comparisons were not based on ERP data time-locked to stimulus onset, but time-locked to the critical boundary position, that is, the offset of the second name (see Bögels et al., 2010; Kerkhofs et al., 2007; Pauker et al., 2011 for comparable analyses). ERP epochs were adjusted to a baseline of 500 ms relative to the offset of the second name. All electrodes except for T7/8 were included in the statistical analysis. Since a preceding study with comparable stimuli (Holzgrefe et al., 2013) yielded a broadly distributed, rather fronto-central CPS without hemispherical bias, anterior, central, and posterior electrodes, respectively, were pooled in three regions consisting of nine electrodes each.

Based on previous literature (e.g., Bögels et al., 2010; Pauker et al., 2011), the time window for the statistical data analysis was defined as 100–400 ms after the offset of the second name. For the statistical analysis of the ERP data, a separate repeated measures analysis of variance (ANOVA) was employed for each experiment. The ANOVA included the factors Condition with four levels (Exp. 1: NO, PI, PIDUR, and PI2DUR, Exp. 2: No, PI2, DUR, and Pi2DUR) and Region (anterior, central, and posterior), participants were entered as a random factor. Where appropriate, a correction according to Greenhouse and Geisser (1959) was applied and corrected p-values are reported. Only statistically significant ($p < .05$) main effects and interactions including the factor Condition were resolved in post hoc comparisons. To adjust for multiple comparisons, p-values of post-hoc paired t-tests were corrected according to Holm (1979). The grand average ERPs displayed in Figures 8.4 and 8.5 were 8 Hz low-pass filtered (Butterworth zero phase filter: high cutoff: 8 Hz; slope: 12 dB/oct) for presentation purposes only, all statistical analyses were applied to unfiltered data.

8.3 Results

8.3.1 Behavioral data

Experiment 1

The total number of missing values in the prosodic judgment task (i.e., no button presses within the 3000 ms response interval) was generally very low. Specifically, 12 out of 6720 trials (0.18%; NO: 1, PI: 5, PIDUR: 5, PI2DUR: 1) were missed in Experiment 1. Thus, a total of 6708 observations was subjected to statistical analysis. In Experiment 1, participants showed the following mean (M) proportions for stimuli judged as containing an IPB at the end of the second name: NO: $M = 0.169$ ($SE = 0.009$), PI: $M = 0.303$ ($SE = 0.011$), PIDUR: $M = 0.638$ ($SE = 0.012$), and PI2DUR: $M = 0.81$ ($SE = 0.01$) (see Figure 8.2).

The statistical analysis revealed a significant main effect of Condition: each of the three conditions with manipulated prosodic cues significantly differed from the baseline condition NO (see Table 8.2 for the statistics of the final mixed logit model), demonstrating a global influence of the acoustic manipulations on boundary perception. Moreover, re-leveling of the final model revealed that all conditions containing manipulated cues differed from each other (PI vs. PIDUR: [$\beta = 1.49$;

SE = 0.08; *z* = 19.42; *p* < .001]; PI vs. PI2DUR: [*β* = 2.43; *SE* = 0.09; *z* = 28.29; *p* < .001]; PIDUR vs. PI2DUR: [*β* = 0.94; *SE* = 0.08; *z* = 11.41; *p* < .001]). Notably, the proportions of trials judged to contain a boundary were significantly below chance for condition NO and PI, but significantly above chance for PIDUR and PI2DUR (all *p* < .001, corrected according to Holm, 1979).

Table 8.2 | Statistical results for the fixed effects of the final mixed logit model of the behavioral data from Experiment 1.

Fixed effects	*β*	*SE*	*z*-Value
Intercept (NO)	-1.69	0.12	-13.95***
PI	0.80	0.08	9.2***
PIDUR	2.29	0.08	26.4***
PI2DUR	3.23	0.09	33.9***

*Note: Significance level: *** p < .001, β = estimate, SE = standard error.*

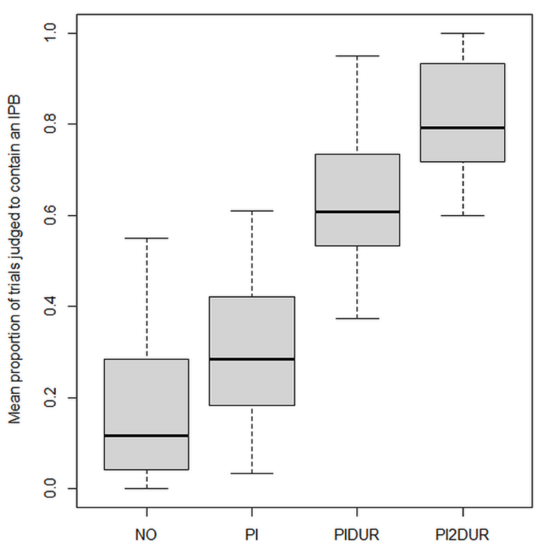

Figure 8.2 | Mean proportions of trials judged as containing an IPB in Experiment 1. Whiskers indicate range of data.

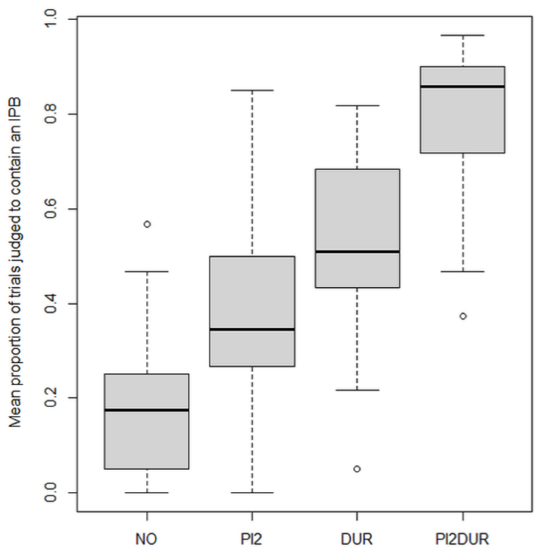

Figure 8.3 | Mean proportions of trials judged as containing an IPB in Experiment 2. Whiskers indicate range of data or 1,5xIQR plus outliers.

Experiment 2

In Experiment 2, 21 out of 7200 trials (0.29%; NO: 10, PI2: 4, DUR: 4, PI2DUR: 3) were missed, yielding a total of 7179 observations that entered the statistical analysis. Participants showed the following mean (*M*) proportions for stimuli judged as containing an IPB at the end of the second name: NO: *M* = 0.175 (*SE* = 0.009), PI2: *M* = 0.365 (*SE* = 0.011), DUR: *M* = 0.538 (*SE* = 0.012), and PI2DUR: *M* = 0.79 (*SE* = 0.01) (see Figure 8.3).

Again, the statistical analysis using mixed logit models revealed a significant main effect of Condition: each of the three conditions with manipulated prosodic cues significantly differed from the baseline condition NO (see Table 8.3). Hence, acoustic cue manipulation generally influenced the perception of the experimental stimuli. However, re-leveling of the final mixed logit model showed that only condition PI2DUR obtained a judgment score that was significantly above the chance level ($p < .001$). In other words, only the combined presentation of pitch and final lengthening cues shifted the participants' decision towards the presence of an IPB. On the contrary, mean proportions of the other conditions were either at chance level (condition DUR, $p = 0.09$) or significantly below chance (condition NO and PI, $p < .001$; all p-values corrected according to Holm, 1979), with the latter indicating that participants predominantly judged these stimuli as not containing a boundary. In addition, the results of model re-leveling showed that all manipulated conditions significantly differed from each other (PI2 vs. DUR: [$\beta = 0.73$; $SE = 0.07$; $z = 10.54$; $p < .001$]; PI2 vs. PI2DUR: [$\beta = 1.98$; $SE = 0.08$; $z = 25.38$; $p < .001$]; DUR vs. PI2DUR: [$\beta = 1.24$; $SE = 0.08$; $z = 16.27$; $p < .001$]).

Table 8.3 | Statistical results for the fixed effects of the final mixed logit model of the behavioral data from Experiment 2.

Fixed effects	β	SE	z-Value
Intercept (NO)	-1.59	0.1	-15.49***
PI2	1.02	0.08	12.71***
DUR	1.75	0.08	22.03***
PI2DUR	2.99	0.09	34.35***

*Note: Significance level: *** p < .001, β = estimate, SE = standard error.*

8.3.2 ERP data

Experiment 1

Figure 8.4 depicts the grand average ERPs from nine representative electrodes time-locked to the end of the second name. Following the end of the second name, the conditions PIDUR and PI2DUR exhibit a positive shift which is more pronounced than in conditions NO and PI.

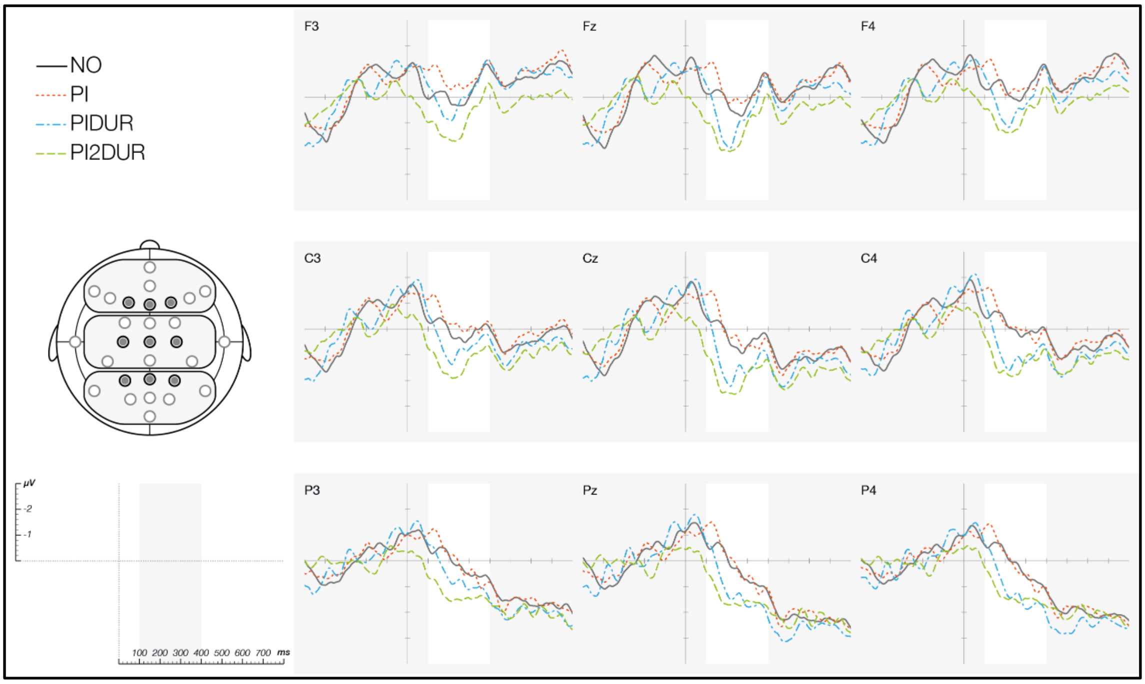

Figure 8.4 | Grand average ERPs from Experiment 1, time-locked to the offset of the second name. Nine representative electrodes are displayed after 8 Hz low-pass filtering. White boxes indicate the time window employed in the statistical analysis (100–400 ms).

A repeated measures ANOVA was computed including the factors Condition (NO, PI, PIDUR, and PI2DUR) and Region (anterior, central, and posterior), participants were entered as a random factor. The fully crossed ANOVA revealed a statistically significant main effect of Condition [$F(3,81) = 19.33$, $MSE = 1.42$, $p < .01$], but no interaction of Condition and Region [$F(6,162) = 1.83$, $MSE = 0.34$] for Experiment 1. Post-hoc paired t-tests revealed that the main effect was driven by the conditions PIDUR and PI2DUR being significantly more positive than condition NO, whereas condition PI did not differ from NO. Likewise, condition PI significantly differed from conditions PIDUR and PI2DUR (all $p < .05$; corrected according to Holm, 1979). Hence, following the end of the second name, a broadly distributed CPS is elicited in response to conditions PIDUR and PI2DUR, whereas no such positivity can be found for condition PI which does not differ from condition NO. Moreover, the conditions PIDUR and PI2DUR were found to differ significantly ($p < .05$; corrected according to Holm, 1979) with condition PI2DUR being even more positive than PIDUR. Accordingly, the CPS found for the combined manipulation of pitch and final lengthening cues is even more pronounced when the global pitch cue is employed.

Experiment 2

Figure 8.5 displays the ERP responses from nine representative electrodes (three from each region of interest). Following the end of the second name, only the condition PI2DUR exhibits a pronounced positive shift, whereas conditions PI2 and DUR overlap with the baseline condition NO. Paralleling the analysis of Experiment 1, a repeated measures ANOVA with the factors Region (anterior, central, and posterior) and Condition (here: NO, PI2, DUR, and PI2DUR) was conducted. For Experiment 2, this revealed a statistically significant main effect of Condition [$F(3,87) = 11.90$, $MSE = 2.10$, $p < .01$], but no interaction of Condition and Region [$F(6,174) = 1.98$, $MSE = 0.35$]. Post-hoc paired t-tests (all $p < .05$; corrected according to Holm, 1979) comparing the four levels of Condition revealed that condition PI2DUR significantly differed from each of the three other conditions, which in turn did not differ from each other. A broadly distributed CPS can thus also be found in Experiment 2, but only for condition PI2DUR. Condition PI2 and DUR, however, did not differ from the baseline condition NO.

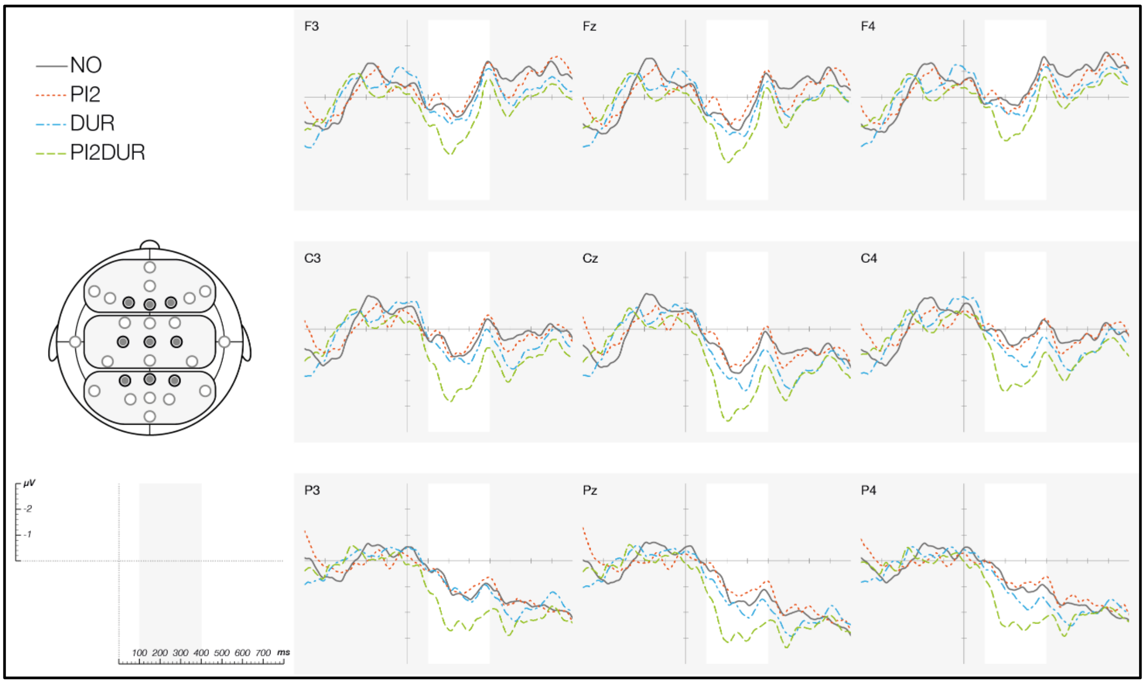

Figure 8.5 | Grand average ERPs from Experiment 2, time-locked to the offset of the second name. Nine representative electrodes are displayed after 8 Hz low-pass filtering. White boxes indicate the time window employed in the statistical analysis (100–400 ms).

Taken together, behavioral and ERP data from both experiments indicate that an IPB is perceived when the two prosodic cues are jointly presented. This holds regardless of the type of pitch manipulation, that is, both the local manipulation of PIDUR and the more global manipulation of PI2DUR were judged to contain an IPB and provoked a CPS. In contrast, such a result was not obtained for final lengthening or pitch as sole cues, once again insensible to the latter being implemented as a local or a global cue. Thus, the single boundary cues examined here were found to be insufficient for IPB perception, both on the electrophysiological and on the behavioral level.

8.4 Discussion

The purpose of the current study was to shed light on the particular contribution of pitch change and final lengthening to German IPB processing. To begin, we found that a combination of pitch and final lengthening cues leads to the perception of an IPB at an otherwise unmarked utterance position. This can be concluded from both behavioral data—demonstrating that participants judged stimuli including both cues predominantly as containing an IPB—and ERP data which revealed that a CPS is elicited if and only if both prosodic cues are presented in the manipulated stimulus material. In addition to underpinning the effectiveness of our experimental design, this result confirms that pause, albeit being an especially salient cue, is not a necessary cue to IPB perception. The finding is in line with previous ERP studies (Steinhauer et al., 1999; Männel, 2009) which demonstrated that a CPS is elicited even if the pause was cut out of the stimulus material. Moreover, the result pattern matches a corpus study on the occurrence and distribution of prosodic phrase boundary cues in German spontaneous speech (Peters et al., 2005) which found pause to be optional and rather infrequent. Although cue combination was in general frequent (61.6% of all boundaries were marked by two or all three cues in combination), Peters et al. (2005) found that pauses only occurred at 38.3% of all boundaries, whereas pitch change (74%) and final lengthening (66.2%) were found to be frequent and thus more reliable cues. Accordingly, the co-occurrence of pitch and lengthening (24.6%) and the combination of all three cues (23.7%) were the most frequent among the cue combinations. The present study suggests that this distribution of prosodic boundary cues in spontaneous speech is driven or at least mutually influenced by their relative importance for perception, especially with regard to pause as a dispensable boundary cue.

Hence, there is converging evidence that the combined presentation of pitch and final lengthening cues is sufficient for IPB perception and to trigger the CPS, notably irrespective of the type of acoustic manipulation (i.e., removing or adding boundary cues). It is yet remarkable that the present study—due to the conjoint use of behavioral and ERP measures—discloses quantitative differences with regard to the combined cues occurrences: Although both combined conditions (i.e., PIDUR and PI2DUR) gave rise to a CPS, a significantly larger positive shift was obtained in response to condition PI2DUR, which contained the rather global, and thus presumably stronger, pitch cue. This finding strongly suggests that CPS amplitude differences are in general suitable to reflect differences in boundary cue strength. In the present case, this difference in cue strength is not based on local (pre-boundary) differences, but is evoked by enhancing the relative effect of the pitch rise through the flattened preceding pitch contour (see, e.g., Clifton et al., 2002; Frazier et al., 2006 for a discussion of non-local boundary effects and global prosodic cue interpretation).

With regard to the impact of single cues, our results consistently show that the presentation of pitch change and final lengthening as sole cues did not shift the listeners' perception towards an IPB interpretation. Instead, stimuli involving only a single cue were either judged as not containing a boundary (condition PI and PI2) or the participants responded at chance level (condition DUR). The behavioral results are mirrored in the ERP data in that no CPS occurred in response to single cue conditions. We obtained this result irrespective of the type of pitch manipulation, that is, the enhanced salience of the condition with a more global pitch cue (PI2) neither improved judgment scores as compared to local boundary tone insertion (PI), nor did it trigger a CPS. Hence, pitch change and final lengthening as single cues are not sufficient for IPB perception. Note that consistent results were also obtained when testing 8-month-old German infants on a comparable set of stimuli using the headturn preference procedure (Wellmann et al., 2012).

At first glance, the result pattern is in contrast to earlier studies on English IPB perception postulating an impact of pitch change and/or final lengthening as sole cues (Aasland & Baum, 2003; Streeter, 1978). A thorough comparison of the study designs, however, reveals that the results have much in common: Streeter (1978) considered pitch change as well as final lengthening to be “effective” cues because she found that the respective acoustic manipulation influenced the participants’ response in that listeners showed a higher tendency to accept the alternative phrasing if at least one of the cues was altered. However, since only the combined manipulation of pitch and final lengthening completely shifted the listeners’ interpretation towards the alternative phrasing, the result pattern is fully in line with the here presented data for German. In the current study, the behavioral data also show that each cue on its own influences the prosodic judgment: In both experiments, significantly larger proportions of trials judged as containing an IPB were obtained for the single cue conditions in comparison to the baseline condition. However, since these larger proportions are still below or at the chance level, we refrain from labeling the single prosodic cues tested here as “effective” in the sense that single cues are sufficient for IPB perception, at least in the context of our investigation. It is noteworthy that the same reasoning holds with regard to cue weighting results from Mandarin Chinese obtained by Yang et al. (2014; see above). They claimed that pitch reset and final lengthening are the weakest boundary cues, implying that the cues are effective on their own since boundary detection rates were higher than for the condition without boundary cues. Crucially, the proportions of detected boundaries were below 60% and the authors do not state whether these detection rates are actually above the chance level. Only pause as a single cue and in combination with final lengthening yielded detection rates that were similar to the natural baseline condition. A combination of pitch and final lengthening was unfortunately not part of the study design. Nonetheless, Yang et al.’s (2014) results are insofar consistent with the current study as (a) pitch change and final lengthening do not differ in boundary cue strength (see also below) and as (b) it seems that neither pitch nor final lengthening as single cues can shift the listeners’ interpretation towards perceiving an IPB, while a combination of two boundary cues does so.

Hence, the evaluation of behavioral and combined electrophysiological measures suggests a methodological refinement: a significant increase in the prosodic judgments towards perceiving a boundary should be distinguished from consistent boundary perception. We found such a significant increase for the single cue conditions in both experiments. However, the increase still leaves the participants opting for a boundary below or at chance level, and it is not accompanied by an appreciable electrophysiological signal of boundary perception. This is distinct, in our results, from the combined effect of final lengthening and pitch manipulation, in which the participants' interpretation was completely shifted in the behavioral task (i.e., IPB perception rates were significantly above chance level) and in which the CPS also indicated an effect of such boundary perception.

Aasland and Baum (2003), in contrast, found that final lengthening as a single cue can shift the listeners' grouping decision towards the alternative prosodic phrasing. There is, yet, an important methodical difference regarding the choice of stimuli and cue size. While we manipulated the stimuli categorically, that is, a cue was either present or absent, Aasland and Baum (2003) chose a 5-step manipulation continuum with a step size of 40 ms. This difference by itself may have played a role in obtaining diverging results. Furthermore, it apparently leads to a larger lengthening effect for the upper end of the continuum than the about 80 ms which were inserted as final lengthening cue in the current study (resulting from applying a lengthening factor of 180% to the final vowel). This cue size was determined by the values that were carefully observed in natural speech, that is, in the stimulus counterpart containing an IPB. Admittedly, we cannot rule out that the outcome of our study may depend on choices made regarding the experimental design, namely cue size and categorical acoustic manipulation. Hence, we can only speculate whether a larger final lengthening cue would have produced a different result. Given that final lengthening as a single cue already influenced the participants' decision in such a way that they responded at chance level, a larger cue size (i.e., more extended lengthening) may have resulted in shifting the response towards IPB interpretation. Notably, the comparison of our results with previous studies does not allow for conclusions to be drawn on a language-specific (i.e., German) cue weighting diverging from other languages. Given the methodological differences explicated above, future research is needed that emphasizes on a cross-linguistic investigation of boundary cue perception by testing several native language populations within the same study design, ideally with identical or at least highly comparable stimuli. To circumvent additional language-specific cues or preferences in the stimulus material, the use of synthetic speech may be another prerequisite for effectively studying the cross-linguistic variation of prosodic cue perception and weighting.

Another limitation of the study design is that apart from the conditions featuring the global pitch manipulation (i.e., PI2 and PI2DUR), non-local prosodic cues that occur distant from the critical IPB position (e.g., on the first name) could not be taken into account. We assume that the current work covers the strongest effects on prosodic boundary processing which are presumably generated by local prosodic cues (i.e., in close vicinity to a boundary position), but our study design clearly does not allow for conclusions to be drawn on the general impact of non-local cues. Explicitly testing the impact of specific non-local prosodic cues (e.g., at prosodic boundaries occurring earlier in the stimulus) on boundary processing—as stated, for instance, in the informative boundary hypothesis (Clifton et al., 2002)—is hence a matter of future research using an adapted experimental design. Importantly, in case of a significant impact of non-local cues we would also expect to elicit a CPS in response to the respective experimental conditions.

This is one of the first ERP studies that systematically investigate the impact of specific boundary cues (for another recent study see Männel & Friederici, 2016). As outlined above, behavioral and ERP data yield in general very similar result patterns. This notable parallelism renders future ERP studies possible, investigating prosodic boundary processing and cue weighting without the need of an explicit metalinguistic task (e.g., judgments or ratings), for example in populations that are not well-suited for judgments (e.g., children; see also Männel & Friederici, 2011; Männel et al., 2013; Männel & Friederici, 2016), during the online processing of syntactic ambiguities or while other tasks are performed. Despite the resemblance of behavioral and ERP data, the combined use allows for a more fine-grained result evaluation: in Experiment 1, the magnitude of the CPS effect discloses a substantial difference between the two combined conditions (PIDUR and PI2DUR) that could not have been inferred from the behavioral results only, given that as all judgment scores generally differed from all other conditions. Further, results of Experiment 2 are of special interest with regard to the cue weighting, that is, the relative importance of each individual boundary cue. Here, an instance of the pitch cue (PI2) and final lengthening were both tested as single cues within one experiment and, therefore, the respective ERPs can be directly compared. Crucially, no significant amplitude or latency differences became manifest in the ERPs elicited for the respective conditions. Hence, although the off-line behavioral results of Experiment 2 are slightly in favor of assuming a stronger impact of final lengthening as compared to pitch change—with a significantly larger mean proportion of trials that were judged to contain a boundary for the first than for the latter—the current ERP results do not suggest a language specific cue weighting with final lengthening or pitch change being a stronger IPB cue than the other one. Instead, pitch change and final lengthening seem to act as interrelated cues that are only effective when occurring in combination.

There is, in fact, growing evidence in the literature for remarkable dependencies in the perception of pitch and duration: Not only that pitch change information has to unfold over time to be perceivable, the perception of (vowel or syllable) duration also seems to depend on f0 characteristics. Yu (2010) and Cumming (2011), for example, showed that participants perceived vowels as longer when the stimuli had a dynamic f0 (as, e.g., in a pitch rise) compared to stimuli with a static f0. Recent work of Brugos and Barnes (2012, 2014b) focused on the auditory kappa effect, which essentially denotes the phenomenon that the perception of time intervals can be distorted by the perception of intervals in pitch space, namely, tonal height. The authors found evidence that the effect does not only occur when perceiving simple tones, but that the perception of pitch changes between monosyllabic words affects the duration perception for the silent intervals between the words. However, a subsequent experiment employed a phrasal grouping task and revealed that the impact of pitch change on prosodic grouping cannot be fully captured by the auditory kappa effect. Instead, pitch information seems to contribute to the perception of prosodic boundaries also independently of segmental or syllable durations. In sum, the interplay of pitch change and final lengthening is not yet well-understood, but a matter of current and future research.

On a final note, the present results do not contradict Gollrad et al.'s (2010) conclusion that durational cues (i.e., pause and final lengthening) have in combination a stronger impact on boundary perception than pitch information. Rather, both studies on German IPB perception provide converging evidence that pitch change as a single boundary cue does not trigger IPB perception.

8.5 Conclusions

The present study was aimed at evaluating the role of pitch change and final lengthening cues in the perception of IPBs. In two experiments we found that a combination of pitch and final lengthening cues is sufficient to shift the listeners' perception towards an IPB interpretation and to elicit a CPS. We therefore conclude that pause, albeit being a very salient cue, is not necessary for IPB perception. Presented as single cues, however, neither pitch change nor final lengthening altered the listeners' predominant judgment or provoked a CPS. Hence, behavioral and ERP data consistently show that the perception of German intonation phrase boundaries is caused by the combination of pitch and final lengthening, but not by pitch or final lengthening alone. The present finding matches data on the occurrence of prosodic boundary cues in German spontaneous speech (Peters et al., 2005) and the prosodic cue perception observed in German 8-month-old infants (Wellmann et al., 2012). Though inferences on language-specific as opposed to cross-linguistic preferences are rather limited, the current results seem to parallel results from others languages. Taken together, there is converging evidence that pitch change and final lengthening are highly interrelated boundary cues playing a decisive role in language comprehension. The combined use of ERPs and prosodic judgments revealed a close relationship between the two measures that should be further explored in future research (e.g., using single-trial ERP analysis), as it may be informative with regard to (language-specific) prosodic cue weighting and possible inter-subject or stimulus-related variability.

Acknowledgements

This research was funded by the German Science Foundation (*Deutsche Forschungsgemeinschaft*, DFG), Priority Program 1234, under Grant FR 2865/2-1 to Isabell Wartenburger and under Grant HO 1960/13- 1 to Barbara Höhle.

9 Infants' processing of prosodic cues: electrophysiological evidence for boundary perception beyond pause detection[3]

Abstract

Infants as young as six months are sensitive to prosodic phrase boundaries marked by three acoustic cues: pitch change, final lengthening, and pause. Behavioral studies suggest that a language-specific weighting of these cues develops during the first year of life; recent work on German revealed that eight-month-olds, unlike six-month-olds, are capable of perceiving a prosodic boundary on the basis of pitch change and final lengthening only. The present study uses Event-Related Potentials (ERPs) to investigate the neuro-cognitive development of prosodic cue perception in German-learning infants. In adults' ERPs, prosodic boundary perception is clearly reflected by the so-called Closure Positive Shift (CPS). To date, there is mixed evidence on whether an infant CPS exists that signals early prosodic cue perception, or whether the CPS emerges only later—the latter implying that infantile brain responses to prosodic boundaries reflect acoustic, low-level pause detection. We presented six- and eight-month-olds with stimuli containing either no boundary cues, only a pitch cue, or a combination of both pitch change and final lengthening. For both age groups, responses to the former two conditions did not differ, while brain responses to prosodic boundaries cued by pitch change and final lengthening showed a positivity that we interpret as a CPS-like infant ERP component. This hints at an early sensitivity to prosodic boundaries that cannot exclusively be based on pause detection. Instead, infants' brain responses indicate an early ability to exploit subtle, relational prosodic cues in speech perception—presumably even earlier than could be concluded from previous behavioral results.

[3] This chapter has been published as:
Holzgrefe-Lang, J., Wellmann, C., Höhle, B., & Wartenburger, I. (2017). Infants' processing of prosodic cues: Electrophysiological evidence for boundary perception beyond pause detection. *Language and Speech*, Sep 1:23830917730590. [Epub ahead of print]

9.1 Introduction

Prosodic phrasing information plays an important role both in language comprehension and in language acquisition due to the coincidence of the edges of major prosodic phrases with syntactic boundaries (e.g., Downing, 1970; Nespor & Vogel, 1986; Selkirk, 2005; Truckenbrodt, 2005). Hence, an influence of prosodic boundary processing on adults' syntactic analysis has been demonstrated by numerous studies (see, for a review, Cutler et al., 1997). Likewise, young infants have been shown to be highly sensitive to prosodic boundary information and to use it for segmentation of clauses or words (see, for a review, Speer & Ito, 2009).

Across languages, a major prosodic boundary, also referred to as an Intonational Phrase boundary or prosodic break (see Bögels et al., 2011), is mainly signaled by three prosodic characteristics: pitch change (i.e., a pitch rise or pitch fall and/or a pitch reset), final lengthening (i.e., an increase in the duration of the segments immediately preceding the boundary), and a pause (e.g., Hirst & Di Cristo, 1998, Nespor & Vogel, 1986; Vaissière, 1983; Wightman et al., 1992).

Although there is a universal tendency to predominantly use these three acoustic cues to mark major prosodic boundaries, the reliance on a particular cue or cue combination when producing or perceiving a prosodic boundary may differ across languages (see, amongst others, Aasland & Baum, 2003; Gollrad et al., 2010; Lehiste, Olive, & Streeter, 1976; Scott, 1982; Streeter, 1978; Yang et al., 2014; Zhang, 2012). In other words, the cue weighting, that is, the relative importance of each cue and combinations of them, varies across languages and may thus be language-specific. For German adults, a recent study (Holzgrefe-Lang et al., 2016) showed that the combination of pitch change and final lengthening without a pause triggers the perception of a major prosodic boundary, while the individual cues alone are not sufficient.

Recently, research on language acquisition has focused on the question of whether, when, and how a language-specific weighting of prosodic cues develops in infancy (e.g., Johnson & Seidl, 2008; Seidl, 2007; Seidl & Cristià, 2008; Wellmann et al., 2012; Wellmann et al., submitted): a core question is whether infants need all three prosodic cues to perceive a major prosodic boundary or whether and at which age a subset is sufficient, presumably reflecting developmental changes towards adult-like boundary perception. Below, the relevant behavioral findings on the development of prosodic boundary perception in English-, Dutch-, and German learning infants are outlined (see also Table 1). Seidl (2007) investigated prosodic cue weighting in six-month-old English-learning infants. Using the headturn preference procedure (HPP), she tested the recognition of previously heard word sequences in longer passages of continuous speech. Infants were familiarized with two prosodic versions of the same word sequence: either the word sequence occurred as a single prosodically well-formed phrase or it spanned two phrases and was prosodically ill-formed. By establishing a preference for passages containing the well-formed word sequences, Seidl (2007) showed that the phrases were recognized even in the absence of final lengthening or the pause but not when the pitch cue was neutralized. Pitch change as a single cue was not sufficient to indicate the perception of a major prosodic boundary. This suggests that by six months, English-learning infants are able to perceive a major prosodic boundary on the basis of a subset of the three main boundary cues, as long as the pitch cue is present. In contrast, an extension of the study with younger English-learning infants (Seidl & Cristià, 2008) revealed that four-month-olds need all three prosodic cues to perceive a major prosodic boundary. Adapting the experimental design to Dutch, Johnson and Seidl (2008) found that pause was a necessary cue for six-month-old Dutch learners. The diverging results across the two languages indicate that infants' prosodic cue weighting is already affected by the ambient language at the age of six months.

Further evidence for an attunement of prosodic boundary cue perception within the first year of life stems from two recent HPP studies with German-learning six- and eight-month-old infants (Wellmann et al., 2012, Wellmann et al., submitted). Here, the stimulus material consisted of short sequences of three coordinated names. The sequences either formed a single major prosodic phrase or they were made up of two phrases separated by a major prosodic boundary. When the boundary was signaled by pitch change and final lengthening, but occurred without a pause, eight-month-olds successfully discriminated the two prosodic patterns, whereas no discrimination was found when it was solely signaled by either pitch change or by final lengthening (Wellmann et al., 2012). In contrast, for German-learning six-month-olds, the combination of pitch and lengthening was not sufficient (Wellmann et al., submitted); they needed a pause cue to discriminate the prosodic patterns. This finding indicates a developmental change in German-learning infants' perception of major prosodic boundaries from rather basic acoustic detection at six months that heavily relies on the pause cue to a more sophisticated linguistic processing of prosodic cues at eight months that mirrors German adults' responses to the same material (Holzgrefe-Lang et al., 2016). Table 9.1 summarizes the behavioral results concerning the perception and weighting of prosodic boundary cues in English-, Dutch-, and German learning infants.

Table 9.1 | Summarized findings regarding infants' prosodic boundary cue perception across languages.

	English-learning 6-month-olds (Seidl, 2007)	**English-learning 4-month-olds** (Seidl & Christià, 2008)	**Dutch-learning 6-month-olds** (Seidl & Johnson, 2008)	**German-learning 8-month-olds** (Wellmann et al., 2012)	**German-learning 6-month-olds** (Wellmann et al., subm.)
Boundary marked by					
– all three cues	**+**	**+**	**+**	**+**	**+**
– pitch change and final lengthening	**+**	**-**	**-**	**+**	**-**
– pitch change and pause	**+**	**-**	**+**	N/A	**+**
– pause and final lengthening	**-**	**-**	N/A	N/A	N/A
– pitch change	**-**	N/A	N/A	**-**	N/A
– final lengthening	**-**	N/A	N/A	**-**	N/A

+: behavioral results indicate infants' boundary perception
-: no indication of boundary perception
N/A: condition was not tested in the respective study
Please note that a direct comparison of results across studies is limited by differences regarding the stimuli and the experimental task (i.e., differentiation vs. clause segmentation).

Event-Related Potential (ERP) research has shown that the processing of a major prosodic boundary is indexed by the so-called closure positive shift (CPS)—a broadly distributed, positive-going component that co-occurs with the closure of a major prosodic boundary (for a review, see Bögels et al., 2011). Crucially, the adult CPS occurs independently of the presence of a pause as a boundary cue (Steinhauer et al., 1999; Männel & Friederici, 2009; Holzgrefe-Lang et al., 2016), which clearly shows that the CPS is not a variant of an early auditory evoked potential that signals the detection of new auditory input after a period of silence and thus does not reflect low-level acoustic processing.

So far, only few attempts have been made to study the electrophysiological correlates of major prosodic boundary perception in infants. Pannekamp et al. (2006) presented eight-month-old German-learning infants with sentences that either contained a sentence-internal, fully marked major prosodic boundary or not. Results revealed a positive shift in response to stimuli containing the major prosodic boundary. Even though this shift was delayed as compared to adults (cf. Pannekamp et al., 2005, Steinhauer et al., 1999), the authors considered it to be an infant CPS reflecting their ability to perceive prosodic boundaries.

A series of recent investigations, however, calls the existence of an infant CPS into question and assigns the emergence of the CPS to the acquisition of sophisticated syntactic knowledge gained within the third year of life. Männel and Friederici (2009) attempted to replicate the infant CPS for five-month-old German-learning infants. Results again showed a positive shift in response to the condition containing an internal prosodic boundary. However, fine-grained data analyses revealed that this positivity could not clearly be attributed to the processing of the major prosodic boundary but seemed to at least partially reflect a low-level response to the new input following the pause (i.e., obligatory onset components, see Kushnerenko et al., 2002). Testing of older children (Männel & Friederici, 2011) indicated that only three- and six-year-olds—but not 21-month-olds—showed an adult-like CPS in response to fully marked major prosodic boundaries. Further, six-year-olds showed an adult-like CPS when exposed to boundaries marked by pitch change and final lengthening only (Männel et al., 2013), while three-year-olds needed the combination of final lengthening and a pause (Männel & Friederici, 2016). This hints at a larger impact of durational cues (i.e., the combination of final lengthening and a pause) on German toddlers' prosodic boundary processing than pitch cues, which is in line with results obtained for German adults (see Gollrad et al., 2010, 2010). It remains however unclear how this electrophysiological development in toddlers and preschoolers relates to findings from behavioral studies that observe a developmental shift in boundary cue weighting within the first year of life (Seidl & Cristià, 2008; Wellmann et al., 2012; Wellmann et al., submitted).

The aim of the present study is hence twofold: first, we intend to find out whether an electrophysiological correlate of infant PB processing (as observed by Pannekamp et al., 2006) can be elicited on the basis of pitch and final lengthening cues only, while no pause boundary cue is present. Second and in the main, we seek to shed light on the development of PB perception within the first year of life from an electrophysiological perspective. We conducted two ERP experiments with six- and eight-month-old infants using the same materials as in the study by Wellmann et al. (2012) that had shown that eight-month-old infants perceived the PB in this material on the basis of pitch and final lengthening cues only. We used only stimuli that did not contain a pause to rule out an impact of obligatory onset components provoked by new acoustic material after a silent pause. Moreover, this bears the advantage that the two rather subtle cues can be investigated irrespective of the impact of the salient pause cue.

If infants already show an electrophysiological indicator for the perception of a PB, it should be elicited in the older age group: Under the assumption that the behavioral results (Wellmann et al., 2012) are mirrored in the ERPs, an electrophysiological correlate should be elicited for eight-month-old infants presented with stimuli containing a combination of pitch and lengthening cues, whereas the brain response to stimuli containing only the pitch cue should not differ from stimuli without PB cues. Moreover, six-month-old infants should not show a specific brain response, given that this age group only recognized PBs containing the pause cue (Wellmann et al., submitted).

9.2 Methods

9.2.1 Participants

In two identical experiments, we tested infants from two age groups, namely six- and eight-month-olds. All infants were from monolingual German-speaking families, born full-term, and normal-hearing. Informed consent was acquired from both parents and the study was approved by the local ethics committee. In the first experiment, 31 six-month-old healthy infants (17 girls) were tested. Their mean age was 6 months, 18 days (range: 6 months 6 days to 6 months 28 days). Fourteen additional infants from this age group were tested but their data was not included in the data analysis for the following reasons: experiment stopped early due to crying or signs of discomfort (6), technical problems (1), and excessive EEG artifacts (7, see section data analysis). In the second experiment, 30 eight-month-old healthy infants (13 girls) were tested. Their mean age was 8 months, 20 days (range: 8 months 3 days to 9 months 6 days). Seventeen further infants from this age group were tested but their data was not included in the data analysis for the following reasons: experiment stopped early due to crying or signs of discomfort (7), excessive EEG artifacts (10).

9.2.2 Stimuli

The stimuli consisted of a sequence of three German names coordinated by und ('and'). The names (*Moni, Lilli,* and *Manu*) contained only sonorant sounds to allow for a reliable measure of the fundamental frequency (f0) and precise acoustic manipulation. For the natural speech recordings, a young female German native speaker was provided with a written stimulus list in which two grouping alternatives were indicated by bracketing as in (10). The speaker was instructed to read the name triples in such a way that the experimenter and an independent second listener, both being naïve to the given bracketing, were able to conceive the bracketing and add it to a plain version of the stimulus list. This ensured that the stimuli were produced with the intended prosodic structure.

(10) a. (Moni und Lilli und Manu)

b. (Moni und Lilli) (und Manu)

Both sequences contained the same string of names and differed only in either grouping all three names together as shown in (10a) or grouping the first two names together and the final one apart as shown in (10b). Thus, sequences of type (10b) contained a PB, whereas sequences of type (10a) did not. Each stimulus recording was preceded by the same context question (Wer kommt? 'Who is coming?') read by the experimenter. The speaker repeated the sequence of each type six times, resulting in six tokens per prosodic type. Recordings were made in an anechoic chamber equipped with an AT4033a audio-technical studio microphone, using a C-Media Wave soundcard at a sampling rate of 22050 Hz with 16 bit resolution.

The acoustic analysis of the recordings revealed clear acoustic differences between the two prosodic phrasings on and after the second name (*Lilli*). In sequences with PB a rise in f0 occurred, starting at the second syllable and leading to a high boundary tone at the final vowel of the second name. This f0 change was 2.5 times greater than the slight rise occurring in sequences without PB. In addition, the duration of the final vowel [i] was 1.8 times longer in the grouping with PB. Finally, a pause with an average duration of 506 ms occurred in sequences with PB, whereas no pause was present in sequences without PB (see Table 9.1). Hence, all acoustic correlates of the three main PB cues were observed in sequences with PB: a change in fundamental frequency (specifically an increased continuation rise, henceforth pitch cue), a lengthening of the final vowel, and the occurrence of a pause.

Table 9.2 | Mean values (range) of the acoustic correlates of prosodic boundary cues.

Prosodic boundary cue	Acoustic correlate	Without PB	With PB
Pitch change (rise and high boundary tone)	f0 rise in Hz	88 (77–110)	220 (197–240)
	Maximum f0 in Hz	277 (264–293)	397 (371–422)
Final lengthening	Final vowel duration in ms	99 (91–110)	175 (162–186)
Pause	Pause duration in ms	—	506 (452–556)

All values pertain to measures on and following the second name in the natural speech recordings.

Given that previous behavioral studies showed that six-month-old English-learners (Seidl, 2007) and eight-month-old German-learning infants are able to perceive a PB on the basis of pitch and final lengthening cues only (Wellmann et al., 2012) we refrained from including natural stimuli with all three boundary cues to avoid an impact of low-level acoustic processing caused by the offset of the pause that may mask ERP effects in response to the other two cues. Instead, the parameters observed in the acoustic analysis of the natural speech recordings served as a reference for the subsequent acoustic manipulations. In particular, mean values pertaining to pitch change and final lengthening (as employed by this speaker) were adopted in generating the experimental stimuli.

The experiment contained stimuli from three prosodic conditions that were systematically constructed from the natural speech sequences without PB by using PRAAT software (Boersma & Weenink, 2010): *Condition NO* served as a baseline condition since no boundary cues were present on the second name (see Figure 9.1A). It contained six original sequences without PB that were scaled to a mean intensity of 70dB. To avoid comparing natural with acoustically manipulated material, the pitch contour of the stimuli was stylized (two semitones). This transformation reduced the number of pitch points and formed the basis for further acoustic manipulations in the other conditions. *Condition PI* consisted of the stylized stimuli from condition NO with an inserted pitch rise on the second name as a sole boundary cue (see Figure 9.1B). To implement the pitch rise, f0 was set to the reference values at the midpoints of the four segments [l], [ɪ], [l], [i], and at the position of the maximum pitch present on the final vowel (see above and Table 9.1). Hereby, a naturally sounding pitch rise of 212 Hz (13.65 semitones) was created. *Condition PIDUR* (see Figure 9.1C) tested the impact of a combination of pitch and final lengthening cues. To this end, a pitch rise was implemented just as described for condition PI and additionally, the final vowel [i] of the second name was lengthened to 180%. This factor was chosen because in the natural stimuli, the crucial vowel was on average 1.8 times longer in sequences with PB than in sequences without PB (see Table 9.1). Following acoustic manipulation, all sequences were resynthesized using the PSOLA function in PRAAT. During the experiment, each stimulus was presented ten times, yielding 60 experimental items per condition and a total of 180 experimental trials.

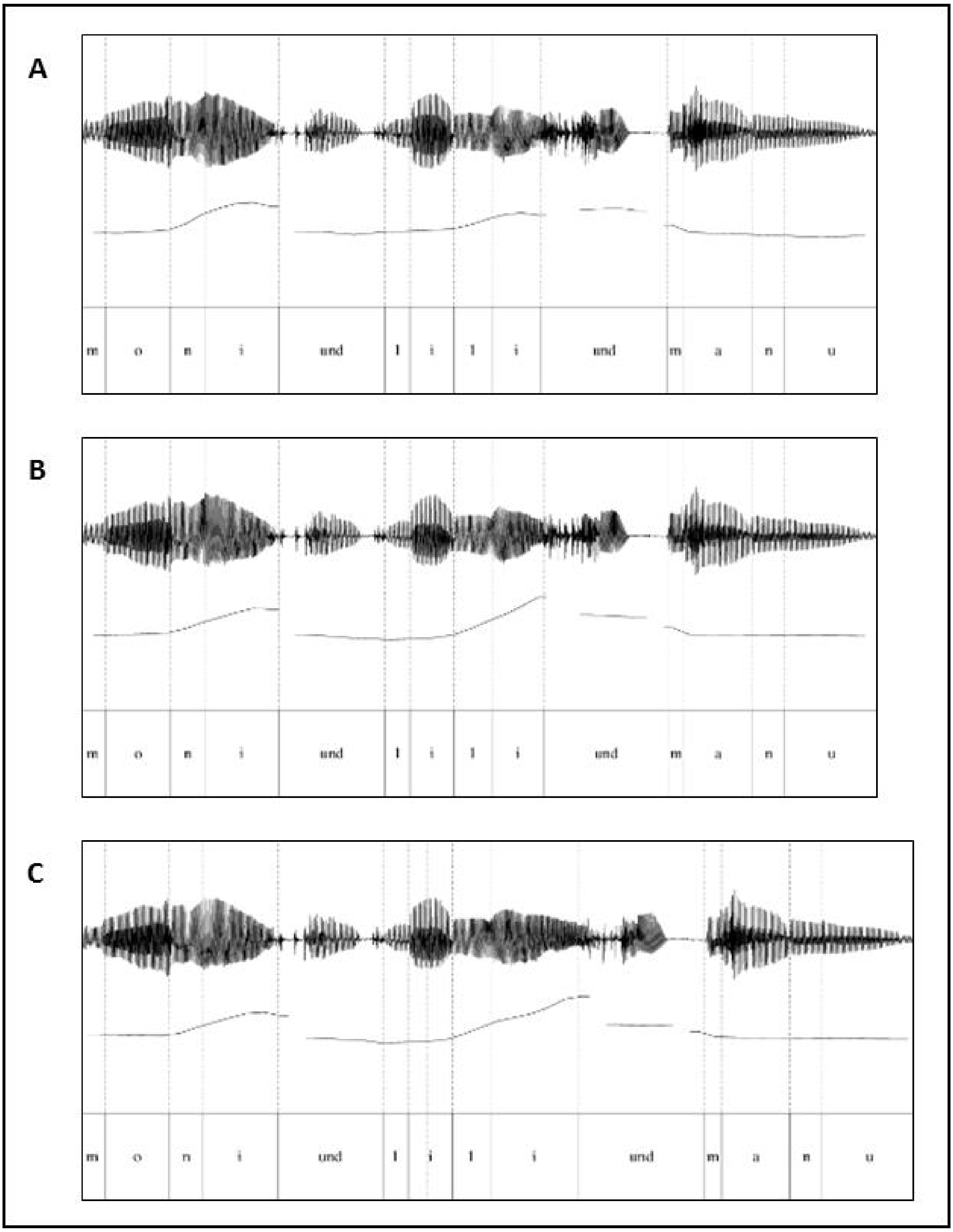

Figure 9.1 | Exemplar waveforms and pitch contours of the experimental stimuli: (A) no-cue condition without boundary cues, (B) pitch-cue condition with a pitch rise on the second name, and (C) combined-cue condition with pitch rise and lengthening of the second name's final vowel.

9.2.3 Procedure

During the EEG recording, infants sat on their parent's lap in an electrically shielded and sound-attenuated booth. Parents perception of the speech played to infants was masked by having them listen to music presented via headphones. Stimulus presentation was controlled by Presentation® software (version 14.1, Neurobehavioral Systems, www.neurobs.com). The 180 experimental stimuli were delivered via two loud speakers and presented in a pseudo-randomized order (i.e., at most two consecutive trials belonged to the same condition and stimuli created from the same base token did not follow each other). To sustain the infants' attention a silent toddler's video was presented temporally unrelated to the acoustic presentation. The total duration of the experiment was approximately 12 minutes. The experiment was interrupted whenever the infant showed any sign of discomfort, and continued only if infant and parent were willing to further participate.

9.2.4 EEG recording and data preprocessing

The EEG was continuously recorded from 31 active Ag/AgCl electrodes (actiCAP, Brain Products, Germany) with a sampling rate of 1000 Hz. The electrodes were placed according to the international 10–10 system (American Clinical Neurophysiology Society, 2006) at the following positions: Fp1/2, Fz, F3/4, F7/8, F9/10, FCz, FC1/2, FC5/6, Cz, C3/4, T7/T8, CPz, CP1/2, CP5/6, Pz, P3/4, P7/8, O1/2. An electrode at AFz served as common ground. Electrical activity was recorded from both mastoids and the EEG recording was referenced on-line to the left mastoid and re-referenced off-line to linked mastoid electrodes. Electrode impedances were in most cases kept below 10 kΩ (at least below 20 kΩ).

The EEG data were off-line processed using Brain Vision Analyzer (version 2.04; Brain Products, Gilching, Germany). A digital band pass filter ranging from 0.3 Hz to 50 Hz was applied to remove slow drifts and muscle artifacts, and we also applied a 50 Hz notch filter.

Segments with a length of 1300 ms, starting 200 ms prior to the onset of the final syllable of the second name, were extracted from the continuous EEG signal. Within these segments, EEG artifacts were detected semi-automatically (rejection criteria: maximal allowed voltage step (gradient): 50 μV, maximal allowed difference in

100 ms intervals: 200 μV, minimal/maximal allowed amplitude: -/+ 200 μV, lowest allowed activity in 100 ms intervals: 0.5 μV). To avoid excessive data loss, segments containing an EEG artifact were not completely removed. Instead, an individual channel mode was applied to allow discarding only the contaminated electrode channels of that segment from further analysis. All electrode channel averages that entered the grand average over participants were based on at least 30 segments (i.e., 50% of all trials) per condition. For the six-month-olds, the mean number of included segments per participant (averaged over electrode sites) was 53 for condition NO (*SD* = 5.29, range 37.6–59.3), 52.5 for condition PI (*SD* = 4.99; range 41.3–59.7), and 53 for condition PIDUR (*SD* = 4.85; 41.7–59.1). For the eight-month-olds, the mean number of included segments per participant (averaged over electrode sites) was 51.6 for condition NO (*SD* = 5.45, range 38.2–58.8), 51.4 for condition PI (*SD* = 5.56; range 38.4–59), and 52.5 for condition PIDUR (*SD* = 4.76; 44.2–58.8). Thus, in both age groups the mean number of included segments did not differ across conditions, nor did it differ between age groups.

9.2.5 Statistical analysis

Statistical comparisons were based on ERP data time-locked to the onset of the final syllable preceding the critical boundary position (see Bögels et al., 2010 for a comparable analysis, and Bögels et al., 2011 for advantages over time-locking to stimulus onset or final syllable offset). ERP epochs were adjusted to a baseline of 200 ms relative to this syllable onset. Note that before this time point, the stimuli from all three conditions were acoustically identical, since acoustic manipulations only affected the final syllable. The time window of the statistical data analysis covered three consecutive 100 ms time windows, ranging from 500 to 800 ms after syllable onset. All electrodes except for Fp1/2 were included in the statistical analysis. Two separate repeated measures analysis of variance (ANOVAs) were employed for the lateral and the midline electrodes. The ANOVA for the lateral electrodes included the factors Condition with three levels (NO, PI, PIDUR), Hemisphere (left, right) and Region (frontal: F3/4, F7/8, F9/10; fronto-central: FC1/2, FC5/6, C3/4; centro-parietal: T7/T8, CP1/2, CP5/6; and posterior: P3/4, P7/8, O1/2). As a result, lateral electrodes were pooled into regions of interest (ROI) containing three electrodes each. The ANOVA for the Midline electrodes included the factors Condition (NO, PI, PIDUR) and Electrode with five levels, corresponding to the single midline electrodes (Fz, FCz, Cz, CPz, Pz). Participants were entered as random factors. Where appropriate, the Huynh-Feldt correction (Huynh & Feldt, 1976) was applied to adjust for sphericity violations and corrected p-values are reported. Only statistically significant ($p < .05$) main effects and interactions including the factor Condition were resolved in post hoc comparisons. The Bonferroni correction was applied to all p-values obtained in post-hoc paired t-tests to adjust for multiple comparisons. The grand average ERPs displayed in Figures 9.2 and 9.3 were 8 Hz low-pass filtered (Butterworth zero phase filter: high cutoff: 8 Hz; slope: 12 dB/oct) for presentation purposes only, all statistical analyses were applied to unfiltered data.

9.3 Results

9.3.1 Experiment 1: six-month-old infants

Figure 9.2 depicts the grand average ERPs of the six-month-old infants. The grand average ERPs are time-locked to the onset of the final syllable preceding the critical boundary position. Especially at frontal and fronto-central electrode sides, the condition PIDUR exhibits a positive shift starting around 300 ms after syllable onset. In contrast, ERP wave forms of condition PI largely resemble the ERPs elicited for the control condition NO.

In the statistical analysis of the lateral electrodes, the repeated measures ANOVA for the time window 500–600 ms revealed a significant main effect of Condition [F (2,60) = 3.91, MSE = 171.28, $p < .05$]. Post-hoc paired t-tests revealed that the main effect of Condition was driven by the condition PIDUR being significantly more positive than the conditions NO and PI (both $p < .05$), whereas condition PI did not differ from NO. For the midline electrodes, the repeated measures ANOVA yielded a significant interaction of Condition and Electrode ([F (8,240) = 2.15, MSE = 13.59, $p < .05$]) for the time window 500–600 ms. This interaction was resolved computing one-way ANOVAs for each electrode, yielding a significant main effect at Fz ([F (2, 60) = 5.3, MSE = 47.19, $p < .05$]). Again, a post-hoc paired t-test showed that the condition PIDUR was significantly more positive than conditions NO and PI (both $p < .05$), which did not differ from each other. Hence, both the analyses of the lateral and the midline electrodes confirm a broadly distributed positivity in response to condition PIDUR, whereas no effect can be found for condition PI that does not differ from condition NO.

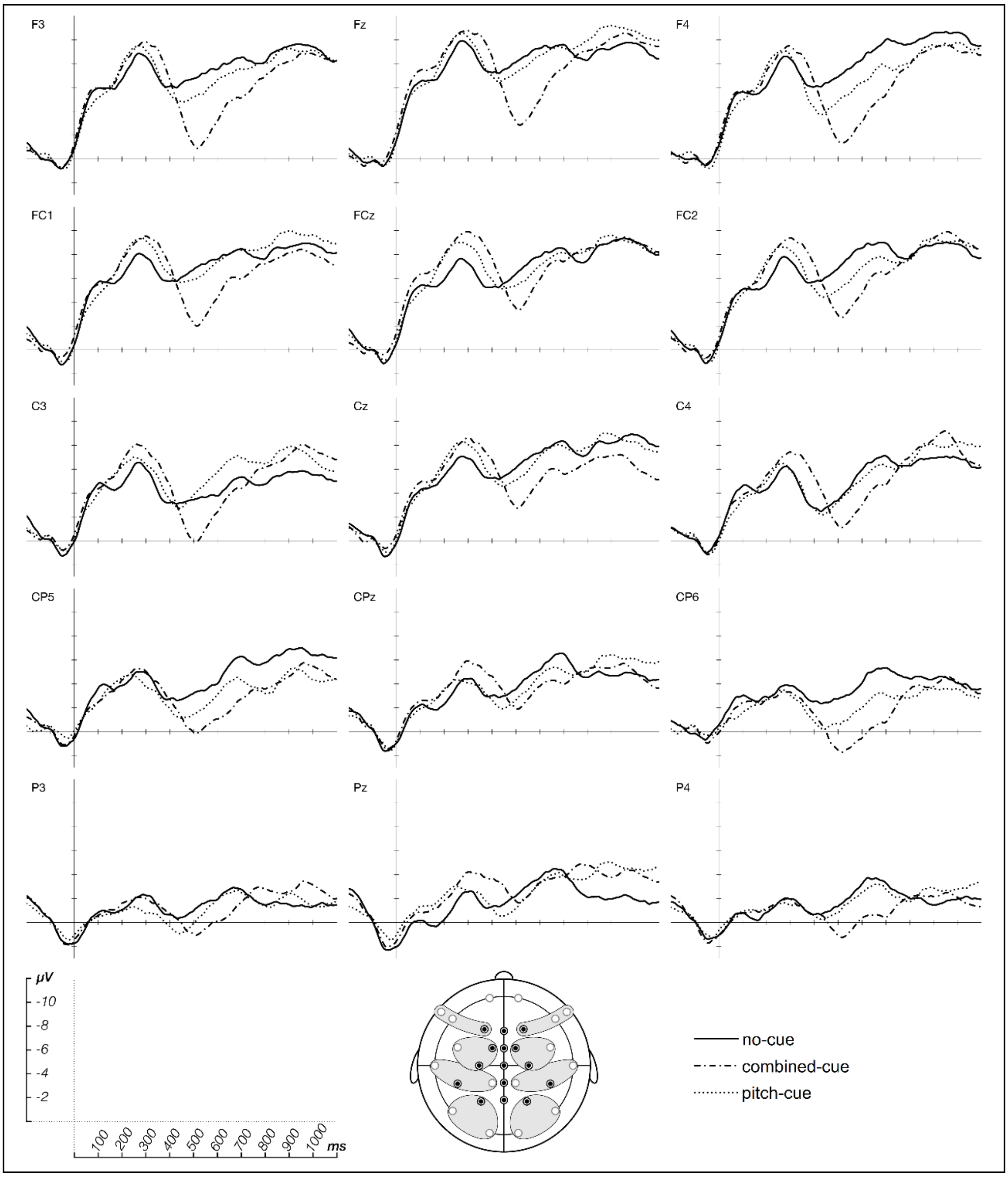

Figure 9.2 | Grand average ERPs from six-month-old infants (Exp. 1), time-locked to the critical final syllable onset. ERPs from 15 representative electrodes are displayed after 8 Hz low-pass filtering.

9.3.2 Experiment 2: eight-month-old infants

Figure 9.3 depicts the grand average ERPs of the eight-month-olds time-locked to the onset of the final syllable. At fronto-central electrode sites, the condition PIDUR exhibits a positive shift that is more pronounced than in the other conditions and starts around 300 ms after syllable onset. Although ERP wave forms of condition PI do rather resemble the ERPs elicited for the control condition NO, visual inspection suggests that there may be a positive deflection for condition PI, too.

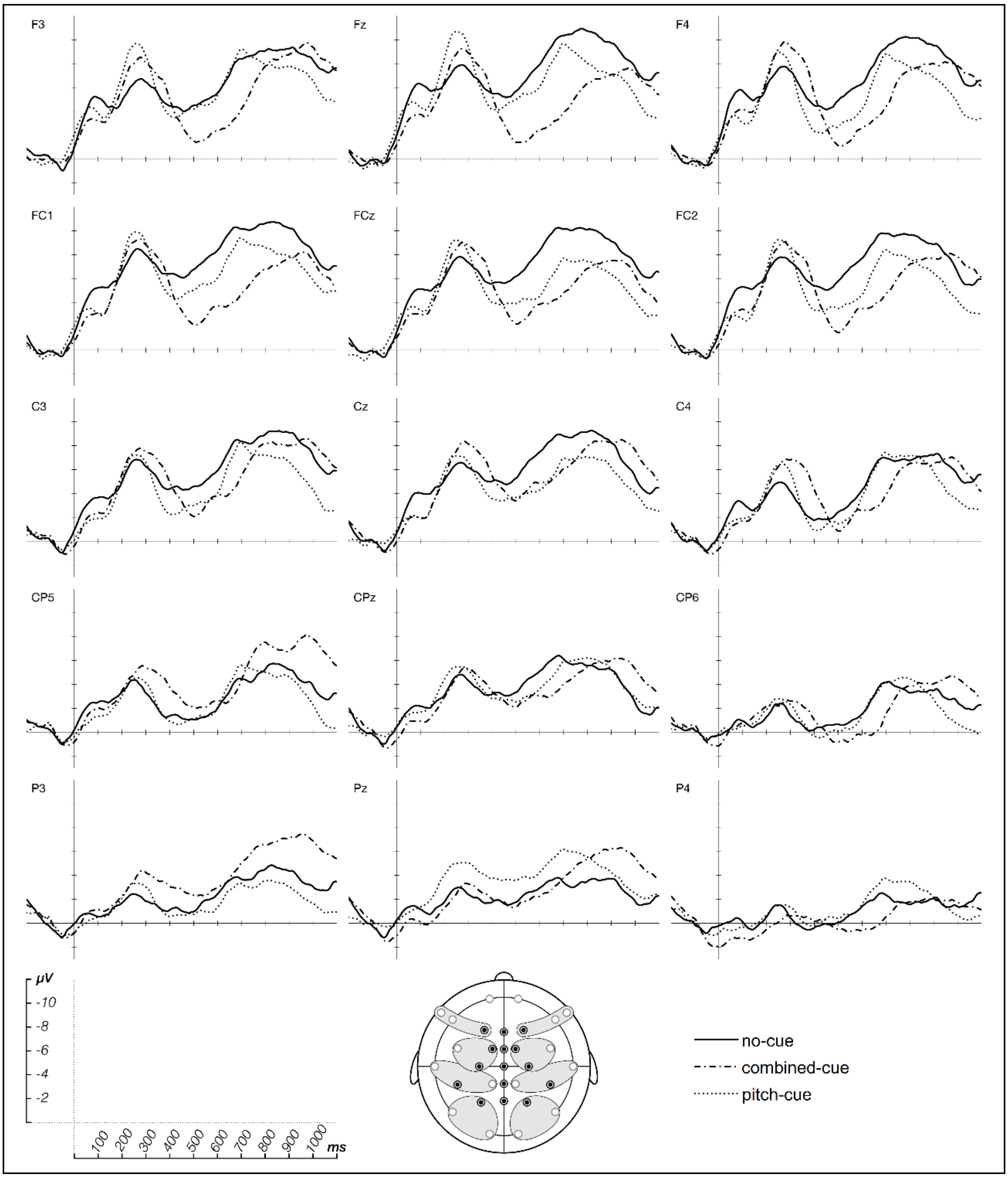

Figure 9.3 | Grand average ERPs from eight-month-old infants (Exp. 2), time-locked to the critical final syllable onset. ERPs from 15 representative electrodes are displayed after 8 Hz low-pass filtering.

The statistical analyses paralleled the analysis of Experiment 1 and revealed significant effects in the time window 600–700 ms: the ANOVA for the lateral electrodes yielded a statistically significant interaction of Condition and Region [F (6,174) = 2.88, MSE = 17.4, $p < .05$] presumably indicating that—as suggested by visual inspection—a possible effect of Condition is larger at frontal than at posterior electrode sites. However, resolving the interaction by computing separate ANOVAs (with the factors Hemisphere and Condition) for each Region only yielded a marginally significant main effect of Condition for the fronto-central region ([F (2, 58) = 3.01, MSE = 40.74, $p = .057$]). The ANOVA of the midline electrodes revealed a significant interaction of Condition and Electrode ([F (8,232) = 2.46, MSE = 25.74, $p < .05$]) in the same time window. Resolving the interaction with one-way ANOVAs for each electrode yielded a significant main effect at Fz ([F (2,58) = 6.41, MSE = 53.9, $p < .01$]). A post-hoc paired t-test comparing the three levels of Condition revealed that PIDUR significantly differed from the two other conditions (both $p < .05$), which in turn did not differ from each other. Hence, there is some evidence, especially from the analysis of the midline electrodes, that a positive shift is elicited in response to condition PIDUR, whereas no such positivity can be confirmed for condition PI. Moreover, in the 700–800 ms time window, again a significant interaction of Condition and Region [F (6,174) = 2.76, MSE = 19.63, $p < .05$] occurred at lateral electrode sites; however, no significant effects were obtained with subsequent smaller ANOVAs for each Region. For the midline electrodes, a significant interaction of Condition and Electrode ([F (8,232) = 2.84, MSE = 25.24, $p < .05$]) was present in the same time window. Resolving the interaction with one-way ANOVAs for each electrode again yielded a significant main effect at Fz ([F (2,58) = 4.14, MSE = 56.7, $p < .05$]). The post-hoc paired t-test comparing the three levels of Condition revealed that PIDUR significantly differed from condition NO ($p < .05$), while condition PI did not differ from the other conditions.

Taken together, statistical analysis confirms that ERP data from both experiments features a positive shift in response to condition PIDUR containing a combination of pitch change and final lengthening cues. In contrast, no statistical difference occurred between condition PI containing the pitch cue only and the control condition NO without boundary cues.

9.4 Discussion

The first aim of this study was to find out whether an electrophysiological correlate of infant PB processing (as observed by Pannekamp et al., 2006) can be elicited on the basis of pitch and final lengthening cues only, while no pause boundary cue is present. In two ERP experiments conducted with six- and eight-month-old infants, we found that the brain response to PBs marked by a combination of pitch change and final lengthening differed from the response to PBs marked by pitch change only and to the control condition without boundary cues. In particular, the grand average ERPs of the six-month-old infants (Experiment 1) display a positive shift in response to the combined cue occurrence (condition PIDUR) that resembles the adult CPS known to mark PB processing in that it has a broad, bi-lateral distribution (see Bögels et al., 2011), though the effect is here most evident at fronto-central electrode sites (see, e.g., Bögels et al., 2010 for a comparable finding). It starts around 300 ms after the onset of the final syllable. Given that the critical final syllable of condition PIDUR is on average 257 ms long, this means that the positivity starts immediately after the offset of the second name, that is, at the boundary position under investigation. This early start and the components' peak around 500 to 600 ms after final syllable onset is in line with results from previous adult studies (e.g., Bögels et al., 2010; Holzgrefe-Lang et al., 2016; Pauker et al., 2011; Peter et al., 2014) using comparable time-locking points (i.e., the final syllable onset or offset) to evaluate CPS effects. The visual observation is confirmed by the statistical analysis yielding significant effects for condition PIDUR, but not for condition PI, as compared to condition NO in the time window 500–600 ms.

In eight-month-olds, a possible CPS effect is less evident in the grand averages. Visual inspection here suggests a positivity in response to condition PIDUR that is less prominent and at some electrode sites (e.g., F4, Cz, see Fig. 3) overlain by a positive deflection in response to condition PI. A reason why, at first sight, the data of the six-month-olds seems to be more meaningful than the grand averages of the older age group could be that the raw data of the eight-month-olds is noisier than the EEG recordings of the six-month-olds. Eight-month-olds were less quiet and acquiescent during the recording session than the younger infants, which might result in noisier data and hence reduced data quality (though this is not reflected in the number of trial loss during artifact rejection in our samples). Nevertheless, statistical analysis revealed significant effects for condition PIDUR in both the 600–700 ms and the 700–800 ms time window, while no such effect could be found in neither time window for condition PI. This turns an alternative explanation unlikely, namely that the visually apparent positive deflection noticed for condition PI may reflect PB processing in stimuli containing only the pitch cue.

In sum, statistical data analysis confirms that six- and eight-month-old infants display a positivity evoked by the combined occurrence of two PB cues, namely pitch change and final lengthening. Importantly, the observed effect cannot be attributed to low-level processes triggered by newly incoming acoustic material after silence, since no pause cue was present. This explanation has previously been used by Männel and Friederici (2009) to account for a CPS-like deflection in eight-month-old infants' brain response to fully marked PBs (Pannekamp et al., 2006). Instead, our results support the idea of a CPS-like infant ERP component as proposed by Pannekamp et al. (2006). However, our results cannot directly be compared to this study, since different experimental stimuli and data analyses were used: While Pannekamp et al. (2006) used full sentences and statistically evaluated differences over large 500 ms time windows time-locked to sentence onset, we used sequences of coordinated names (see Example 10) and specifically examined fine-grained differences at the PB position by time-locking the grand average ERPs to the preceding syllable. The same limitation holds when comparing our results to the series of studies presented by Männel and colleagues (Männel & Friederici, 2009, 2011; Männel et al., 2013). Still, our results conflict with their findings in that a positivity obtained for five- and 21-month-old German learners was attributed to low-level processes due to the onset of the phrase following a PB with pause (Männel & Friederici, 2009, 2011), whereas an adult-like CPS was only encountered for children from three years of age on. The authors link the CPS to complex syntactic-prosodic structure perception, enhanced through syntactic development within the third year of life. Moreover, even in toddlers these mechanisms have been supposed to require the salient pause cue, since a CPS in response to stimuli lacking the pause cue was not found for three-year-olds, but only for six-year-olds (Männel et al., 2013). We suppose that the carefully evaluated effects reported by Männel and colleagues, that is, a broad positive shift only present from the age of three years onwards, may indeed represent adult-like structure perception, which may not or not exclusively be prosody-driven, but influenced by complex syntactic knowledge or structure processing expectations. However, as behavioral findings across languages (Seidl, 2007; Wellmann et al., 2012) show consistent evidence for successful PB perception devoid the pause cue in the second half of the first year of life, it is likely that this early ability is also reflected in the infants' brain responses. Moreover, there is ample behavioral evidence (e.g., Gout et al., 2004; Nazzi et al., 2000; Seidl, 2007; Soderstrom et al., 2003; Shukla et al., 2011) that infants use PB cues to segment the speech input and to extract meaningful units and thus process them linguistically. Therefore, based on our findings we suggest that already in infants a small CPS signals PB processing driven by the prosodic cues provided in the speech stream. This infant CPS presumably reflects the infants' sensitivity to PB cues and the acquisition of prosodic cue processing abilities related to the ambient language.

Regarding our second aim, that is, to characterize the development of infant PB processing, we thus conclude that both six- and eight-month-old German learning infants process a PB marked by a combination of pitch change and final lengthening cues, while the pitch cue alone is insufficient in both age groups. The finding that pitch change presented as a single cue did not trigger boundary perception was expected, as this matches behavioral results of eight-month-olds (Wellmann et al., 2012) failing to detect a PB signaled by the pitch cue only. Moreover, behavioral and electrophysiological responses to comparable stimuli (Holzgrefe-Lang et al., 2016) yielded no evidence for PB perception provoked by the single pitch cue in adults.

With regard to the combined cue occurrence, however, a development from six to eight months was hypothesized since in behavioral experiments (Wellmann et al., 2012; Wellmann et al., submitted) the younger age group needed the pause cue while the eight-month-olds perceived a PB on the basis of pitch change and final lengthening only. This developmental shift is not reflected in the present ERP results, as six-month-olds already show a positivity despite the absence of the pause cue. There is, however, ample evidence from developmental studies that a specific brain response may precede the corresponding behavioral response in the course of development, given that ERPs are measured on-line, and hence, do not depend on task demands or an overt response performance (see Männel & Friederici, 2008). On the contrary, the HPP as an off-line method requires the performance of a task which involves several processing steps until a (motor) response can be measured. Specifically, in a HPP experiment with a familiarization phase as conducted by Wellmann et al. (submitted), infants first need to listen and to memorize one speech stimulus. Then, at test, this stimulus needs to be recognized and discriminated from another, newly presented stimulus condition. Moreover, infants' discrimination is measured only indirectly, namely via orientation times towards the side of presentation. Therefore, different cognitive demands at test may explain the asymmetry between the behavioral and the electrophysiological level in six-month-olds' processing of PBs without a pause. For instance, the recognition of the ambient language's dominant stress pattern has been shown for four-month-old German learners using ERPs (Friederici et al., 2007) while the corresponding behavioral preference is evident only by six months (Höhle et al., 2009).

Given that a certain acquisition step may first be evident in ERPs, and, more indirectly assessed, only later in a behavioral response, we can only speculate that the developmental shift in German infants' PB perception from rather basic acoustic detection to a more sophisticated linguistic processing of PB cues may be evidenced earlier on the electrophysiological level. In other words, it is reasonable that the same result pattern (i.e., a younger age group relies heavily on the pause cue while older infants perceive a PB on the basis of pitch change and final lengthening only) occurs when comparing ERPs of even younger, for instance four-month-old infants with the data obtained for the six-month-olds.

Although the current ERP results do not manifest a developmental decline in the reliance on the pause cue, they still shed light on the infant learner's ability to make use of rather subtle, gradient prosodic cues. Pitch and final lengthening constitute relational information in that they can only function as PB cues when they are processed with respect to global prosodic information to recognize changes in pitch and duration, whereas pause is categorical information that can be recognized locally. While there is much evidence for infants' particular sensitivity to pauses (e.g., Hirsh-Pasek et al., 1987; Morrongiello & Trehub, 1987; Schmitz, 2008), the current finding that German learning infants consistently perceive a PB despite the lack of a pause additionally hints at an early adult-like usage of relational prosodic cues. The emergence of this ability presumably depends on the role of these cues in the ambient language, which is for instance reflected in a larger reliance on the pause of Dutch six-month-olds (Johnson & Seidl, 2008) as compared to their English learning age mates (Seidl, 2007). Future cross-linguistic ERP research is needed to consolidate the concept of an infant CPS reflecting a prosodic cue processing development that is specific to the ambient language. This may reveal cue weighting differences and shed light on the underlying processes leading from very early signal detection to adult-like, expectation-based and structure-driven prosody processing.

9.5 Conclusions

The present results suggest that an infant ERP component resembling the adult CPS signals PB perception. Precisely, both six- and eight-month-olds showed a broadly distributed positive deflection in response to PBs marked by pitch change and final lengthening only, but devoid of a pause. Therefore, the observed effect cannot be assigned to low-level processes triggered by newly incoming acoustic material after silence. Rather, it indicates that both six- and eight-month-old infants perceive a PB when it is signaled by the co-occurrence of pitch change and final lengthening, while the single pitch cue is insufficient for PB perception in both age groups. Hence, we conclude that on the electrophysiological level, infants as young as six months display an ability to make use of a combination of relational PB cues, namely pitch change and final lengthening. Our findings indicate an early overcoming of the presumed learner's strategy to heavily rely on pause occurrences for phrase boundary processing and speech segmentation.

Acknowledgements

This research was funded by a grant from the German Science Foundation (*Deutsche Forschungsgemeinschaft*, DFG), priority program SPP 1234, to Isabell Wartenburger and Barbara Höhle (FR 2865/2-1; HO 1960/13-1). We are grateful to the participating families for their commitment to our research. Furthermore, we would like to thank Romy Räling for stimuli production, and Laura Hahn, Christiane Patzwald, and Susanne Pelke for assistance in data acquisition.

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*(3), 385–399. https://doi.org/10.1016/S0093-934X(03)00138-X

American Clinical Neurophysiology Society. (2006). Guideline 5: Guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology*, *23*(2), 107–110.

Baayen, R. H. (2008). *Analysing linguistic data.* Cambridge: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Bates, D. M., Maechler, M., & Dai, B. (2009). Lme4: Linear-mixed-effect models using S4. Vienna, Austria: R Foundation for Statistical Computing.

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*(6), 644–663. https://doi.org/10.1016/0749-596X(91)90030-N

Beckman, M. E., & Pierrehumbert, J. B. (1986). Intonational structure in Japanese and English. *Phonology Yearbook*, *3*, 255–310. https://doi.org/10.1017/S095267570000066X

Boersma, P., & Weenink, D. (2010). Praat: doing phonetics by computer (version 5.1.29) [Computer program]. Retrieved from http://www.praat.org

Bögels, S., Schriefers, H., Vonk, W., & Chwilla, D. J. (2011). Prosodic Breaks in Sentence Processing Investigated by Event-Related Potentials. *Language and Linguistics Compass*, *5*(7), 424–440. https://doi.org/10.1111/j.1749-818x.2011.00291.x

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*(5), 1036–1053. https://doi.org/10.1162/jocn.2009.21269

Brown, M., Salverda, A. P., Dilley, L. C., & Tanenhaus, M. K. (2011). Expectations from preceding prosody influence segmentation in online sentence processing. *Psychonomic bulletin & review*, *18*(6), 1189–1196. https://doi.org/10.3758/s13423-011-0167-9

Brugos, A., & Barnes, J. (2012). The auditory kappa effect in a speech context. In Qiuwu Ma, Hongwei Ding, Daniel Hirst (Ed.), *Proceedings of 6th International Conference on Speech Prosody* (pp. 151–154). Shanghai: Tongji University Press.

Brugos, A., & Barnes, J. (2014a). Dynamic pitch and pitch range interact in distortions of perceived duration of American English speech tokens. *The Journal of the Acoustical Society of America*, *135*(4), 2425. https://doi.org/10.1121/1.4878063

Brugos, A., & Barnes, J. (2014b). Effects of dynamic pitch and relative scaling on the perception of duration and prosodic grouping in American English. In Campbell, Gibbon, & Hirst (Eds.), *Proceedings of Speech Prosody 7* (pp. 388–392). Dublin, Ireland.

Carlson, K., Clifton, C., & Frazier, L. (2001). Prosodic Boundaries in Adjunct Attachment. *Journal of Memory and Language*, *45*(1), 58–81. https://doi.org/10.1006/jmla.2000.2762

Clifton, C., Carlson, K., & Frazier, L. (2002). Informative Prosodic Boundaries. *Language and Speech*, *45*(2), 87–114. https://doi.org/10.1177/00238309020450020101

Clifton, C., Carlson, K., & Frazier, L. (2006). Tracking the what and why of speakers' choices: Prosodic boundaries and the length of constituents. *Psychonomic bulletin & review*, *13*(5), 854–861. https://doi.org/10.3758/BF03194009

Cole, J. (2014). Prosody in context: A review. *Language, Cognition and Neuroscience*, *30*(1-2), 1–31. https://doi.org/10.1080/23273798.2014.963130

Cumming, R. E. (2010). The interdependence of tonal and durational cues in the perception of rhythmic groups. *Phonetica*, *67*(4), 219–242. https://doi.org/10.1159/000324132

Cumming, R. (2011). The effect of dynamic fundamental frequency on the perception of duration. *Journal of Phonetics*, *39*(3), 375–387. https://doi.org/10.1016/j.wocn.2011.01.004

Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, *40*(2), 141–201.

Downing, B. T. (1970). *Syntactic Structure and Phonological Phrasing in English* (Doctoral dissertation). University of Texas, Austin.

Drury, J. E., Baum, S. R., Valeriote, H., & Steinhauer, K. (2016). Punctuation and Implicit Prosody in Silent Reading: An ERP Study Investigating English Garden-Path Sentences. *Frontiers in Psychology*, *7*, 1375. https://doi.org/10.3389/fpsyg.2016.01375

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*(10), 1696–1711. https://doi.org/10.1162/jocn.2006.18.10.1696

Féry, C., & Truckenbrodt, H. (2005). Sisterhood and tonal scaling. *Studia Linguistica*, *59*(2-3), 223–243. https://doi.org/10.1111/j.1467-9582.2005.00127.x

Fodor, J. D. (1998). Learning to parse? *Journal of Psycholinguistic Research*, *27*(2), 285–319. https://doi.org/10.1023/A:1023258301588

Frazier, L., Carlson, K., & Clifton, C. (2006). Prosodic phrasing is central to language comprehension. *Trends in Cognitive Sciences*, *10*(6), 244–249. https://doi.org/10.1016/j.tics.2006.04.002

Friederici, A. D., Friedrich, M., & Christophe, A. (2007). Brain responses in 4-month-old infants are already language specific. *Current Biology*, *17*(14), 1208–1211. https://doi.org/10.1016/j.cub.2007.06.011

Gee, J. P., & Grosjean, F. (1983). Performance structures: A psycholinguistic and linguistic appraisal. *Cognitive Psychology*, *15*(4), 411–458. https://doi.org/10.1016/0010-0285(83)90014-2

Gervain, J., & Mehler, J. (2010). Speech perception and language acquisition in the first year of life. *Annual review of psychology*, *61*, 191–218. https://doi.org/10.1146/annurev.psych.093008.100408

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

Goldman-Eisler, F. (1968). *Psycholinguistics experiments in spontaneous speech*. London, New York: Academic Press.

Gollrad, A. (2013). *Prosodic cue weighting in sentence comprehension: Processing German case ambiguous structures* (Doctoral dissertation). Universität Potsdam.

Gollrad, A., Sommerfeld, E., & Kügler, F. (2010). Prosodic cue weighting in disambiguation: Case ambiguity in German. In M. Hasegawa-Johnson (Ed.), *Proceedings of Speech Prosody 5* (Vol. 100165, pp. 1–4). Chicago, IL.

Gout, A., Christophe, A., & Morgan, J. L. (2004). Phonological phrase boundaries constrain lexical access II. Infant data. *Journal of Memory and Language*, *51*(4), 548–567. https://doi.org/10.1016/j.jml.2004.07.002

Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and clinical neurophysiology*, *55*(4), 468–484.

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*, 95–112.

Gussenhoven, C. (1992). Sentence accents and argument structure. In I. Roca (Ed.), *Thematic Structure: Its Role in Grammar* (pp. 79–106). Berlin: Foris.

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*(3), 269–286.

Hirst, D., & Di Cristo, A. (1998). A Survey of Intonation Systems. In D. Hirst & A. Di Cristo (Eds.), *Intonation Systems A Survey of Twenty Languages* (pp. 1–44). Cambridge: Cambridge University Press.

Höhle, B. (2009). Bootstrapping mechanisms in first language acquisition. *Linguistics*, *47*(2), 359–382. https://doi.org/10.1515/LING.2009.013

Höhle, B., Bijeljac-Babic, R., Herold, B., Weissenborn, J., & Nazzi, T. (2009). Language specific prosodic preferences during the first half year of life: evidence from German and French infants. *Infant Behavior and Development*, *32*(3), 262–274. https://doi.org/10.1016/infbeh.2009.03.004

Holm, S. (1979). A simple sequential rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*, 65–70.

Holzgrefe, J., Wellmann, C., Petrone, C., Truckenbrodt, H., Höhle, B., & Wartenburger, I. (2013). Brain response to prosodic boundary cues depends on boundary position. *Frontiers in Psychology*, *4*, 421. https://doi.org/10.3389/fpsyg.2013.00421

Holzgrefe-Lang, J., Wellmann, C., Petrone, C., Räling, R., Truckenbrodt, H., Höhle, B., & Wartenburger, I. (2016). How pitch change and final lengthening cue boundary perception in German: converging evidence from ERPs and prosodic judgements. *Language, Cognition and Neuroscience*, *31*(7), 904–920. https://doi.org/10.1080/23273798.2016.1157195

Huynh, H., & Feldt, L. S. (1976). Estimation of the Box Correction for Degrees of Freedom from Sample Data in Randomized Block and Split-Plot Designs. *Journal of Educational Statistics*, *1*(1), 69–82. https://doi.org/10.3102/10769986001001069

Hwang, H., & Schafer, A. J. (2009). Constituent length affects prosody and processing for a dative NP ambiguity in Korean. *Journal of Psycholinguistic Research*, *38*(2), 151–175. https://doi.org/10.1007/s10936-008-9091-1

Hwang, H., & Steinhauer, K. (2011). Phrase Length Matters: The Interplay between Implicit Prosody and Syntax in Korean "Garden Path" Sentences. *Journal of Cognitive Neuroscience*, *23*(11), 3555–3575. https://doi.org/10.1162/jocn_a_00001

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*(1), 8–13. https://doi.org/10.1097/WNR.0b013e328330251d

Jaeger, T. F. (2008). Categorical Data Analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. *Journal of Memory and Language*, *59*(4), 434–446. https://doi.org/10.1016/j.jml.2007.11.007

Johnson, E. K., & Seidl, A. (2008). Clause segmentation by 6-month-old infants: A Crosslinguistic perspective. *Infancy*, *13*(5), 440–455. https://doi.org/10.1080/15250000802329321

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*(2), 252–293.

Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, *29*, 1–23.

Kentner, G., & Féry, C. (2013). A new approach to prosodic grouping. *The Linguistic Review*, *30*(2), 277–311. https://doi.org/10.1515/tlr-2013-0009

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*(9), 1421–1434. https://doi.org/10.1162/jocn.2007.19.9.1421

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. https://doi.org/10.1016/j.brainres.2008.05.034

Kjelgaard, M. M., & Speer, S. R. (1999). Prosodic Facilitation and Interference in the Resolution of Temporary Syntactic Closure Ambiguity. *Journal of Memory and Language*, *40*, 153–194.

Kohler, K. J. (1983). Prosodic Boundary Signals in German. *Phonetica*, *40*(2), 89–134. https://doi.org/10.1159/000261685

Kushnerenko, E., Ceponiene, R., Balan, P., Fellman, V., Huotilaine, M., & Näätäne, R. (2002). Maturation of the auditory event-related potentials during the first year of life. *NeuroReport*, *13*(1), 47–51.

Ladd, D. R. (1988). Declination ‘‘reset’’ and the hierarchical organization of utterances. *The Journal of the Acoustical Society of America*, *84*(2), 530. https://doi.org/10.1121/1.396830

Lehiste, I., Olive, J. P., & Streeter, L. A. (1976). Role of duration in disambiguating syntactically ambiguous sentences. *Journal of the Acoustical Society of America*, *60*(5), 1199–1202. https://doi.org/10.1121/1.381180

Lehiste, I. (1973). Phonetic Disambiguation of Syntactic Ambiguity. *The Journal of the Acoustical Society of America*, *53*(1), 380. https://doi.org/10.1121/1.1982702

Lehiste, I. (1975). Influence of F0 pattern on the perception of duration. *The Journal of the Acoustical Society of America*, *58*(S1), S91. https://doi.org/10.1121/1.2002390

Lehnert-LeHouillier, H. (2007). The influence of dynamic f0 on the perception of vowel duration: cross-linguistic evidence. In J. Trouvain & W. J. Barry (Eds.), *Proceedings of ICPhS XVI* (pp. 757–760). Dudweiler: Pirrot.

Li, W., & Yang, Y. (2009). Perception of prosodic hierarchical boundaries in Mandarin Chinese sentences. *Neuroscience*, *158*(4), 1416–1425. https://doi.org/10.1016/j.neuroscience.2008.10.065

Luck, S. J. (2012). Event-related potentials. In H. M. Cooper (Ed.), *APA handbooks in psychology. APA handbook of research methods in psychology.* Washington, DC: American Psychological Association. https://doi.org/10.1037/13619-028

Luck, S. J. (2014). *An introduction to the event-related potential technique* (Second edition). *A Bradford book*. Cambridge, Massachusetts, London, England: The MIT Press.

Luck, S. J., & Kappenman, E. S. (Eds.). (2012). *The Oxford handbook of event-related potential components*. *Oxford library of psychology*. Oxford: Oxford Univ. Press.

Männel, C. (2009). *Prosodic processing during language acquisition: electrophysiological studies on intonational phrase processing*. *MPI Series in Human Cognitive and Brain Sciences: Vol. 115*. Leipzig: Max Planck Institute for Human Cognitive and Brain Sciences.

Männel, C., & Friederici, A. D. (2008). Event-related brain potentials as a window to children's language processing. In I. A. Sekerina, E. M. Fernández, & H. Clahsen (Eds.), *Language Acquisition and Language Disorders: Vol. 44. Developmental psycholinguistics. On-line methods in children's language processing* (Vol. 44, pp. 29–72). Amsterdam: Benjamins. https://doi.org/10.1075/lald.44.04man

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*(10), 1988–2006. https://doi.org/10.1162/jocn.2009.21221

Männel, C., & Friederici, A. D. (2011). Intonational phrase structure processing at different stages of syntax acquisition: ERP studies in 2-, 3-, and 6-year-old children. *Developmental Science*, *14*(4), 786–798. https://doi.org/10.1111/j.1467-7687.2010.01025.x

Männel, C., & Friederici, A. D. (2016). Neural correlates of prosodic boundary perception in German preschoolers: If pause is present, pitch can go. *Brain Research*, *1632*, 27–33. https://doi.org/10.1016/j.brainres.2015.12.009

Männel, C., Schipke, C. S., & Friederici, A. D. (2013). The role of pause as a prosodic boundary marker: language ERP studies in German 3- and 6-year-olds. *Developmental Cognitive Neuroscience*, *5*, 86–94. https://doi.org/10.1016/j.dcn.2013.01.003

Marcus, M., & Hindle, D. (1990). Description theory and intonation boundaries. In G. Altman (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 483–512). Cambridge, MA: MIT Press.

Martin, J. G. (1970). On judging pauses in spontaneous speech. *Journal of Verbal Learning and Verbal Behavior*, *9*(1), 75–78. https://doi.org/10.1016/S0022-5371(70)80010-X

Morgan, J. L., & Demuth, K. (Eds.). (1996). *Signal to syntax: bootstrapping from speech to grammar in early acquisition*. Mahwah, N.J.: Lawrence Erlbaum.

Morrongiello, B. A., & Trehub, S. E. (1987). Age-related changes in auditory temporal perception. *Journal of Experimental Child Psychology*, *44*(3), 413–426.

Nakatani, C. H., & Hirschberg, J. (1994). A corpus-based study of repair cues in spontaneous speech. *The Journal of the Acoustical Society of America*, *95*(3), 1603–1616. https://doi.org/10.1121/1.408547

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*(1), 123–147. https://doi.org/10.1207/S15327078IN0101_11

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*(3), 407–421. https://doi/10.1162/0898929053279450

Pannekamp, A., Weber, C., & Friederici, A. D. (2006). Prosodic processing at the sentence level in infants. *NeuroReport*, *17*(6), 675–678.

Pauker, E., Itzhak, I., Baum, S. R., & Steinhauer, K. (2011). Effects of cooperating and conflicting prosody in spoken english garden path sentences: ERP evidence for the boundary deletion hypothesis. *Journal of Cognitive Neuroscience*, *23*(10), 2431–2451. https://doi.org/10.1162/jocn.2011.21610

Peppe, S., Maxim, J., & Wells, B. (2000). Prosodic Variation in Southern British English. *Language and Speech*, *43*(3), 309–334. https://doi.org/10.1177/00238309000430030501

Peter, V., McArthur, G., & Crain, S. (2014). Using event-related potentials to measure phrase boundary perception in English. *BMC Neuroscience*, *15*, 129. https://doi.org/10.1186/s12868-014-0129-z

Peters, B. (2005). Weiterführende Untersuchungen zu prosodischen Grenzen in deutscher Spontansprache. In K. J. Kohler, F. Kleber, & B. Peters (Eds.), *AIPUK. Prosodic structures in German spontaneous speech (AIPUK 35a)* (pp. 203–345). Kiel: IPDS.

Peters, B., Kohler, K. J., & Wesener, T. (2005). Phonetische Merkmale prosodischer Phrasierung in deutscher Spontansprache. In K. J. Kohler, F. Kleber, & B. Peters (Eds.), *AIPUK. Prosodic structures in German spontaneous speech (AIPUK 35a)* (pp. 143–184). Kiel: IPDS.

Petrone, C., Truckenbrodt, H., Wellmann, C., Holzgrefe-Lang, J., Wartenburger, I., & Höhle, B. (in press). Prosodic boundary cues in German: Evidence from the production and perception of bracketed lists. *Journal of Phonetics*.

Picton, T. W., Hillyard, S. A., Krausz, H. I., & Galambos, R. (1974). Human auditory evoked potentials. I. Evaluation of components. *Electroencephalography and clinical neurophysiology*, *36*(2), 179–190.

Pisoni, D. B. (1976). Fundamental frequency and perceived vowel duration. *The Journal of the Acoustical Society of America*, *59*(S1), S39-S39. https://doi.org/10.1121/1.2002669

Price, P. J., Ostendorf, M., Shattuck-Hufnagel, S., & Fong, C. (1991). The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America*, *90*(6), 2956–2970.

R Core Team. (2013). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Roll, M., & Horne, M. (2011). Interaction of right- and left-edge prosodic boundaries in syntactic parsing. *Brain Research*, *1402*, 93–100. https://doi.org/10.1016/j.brainres.2011.06.002

Sanderman, A. A., & Collier, R. (1997). Prosodic phrasing and comprehension. *Language and Speech*, *40*(4), 391–409.

Schafer, A. J. (1997). *Prosodic parsing: The role of prosody in sentence comprehension* (Doctoral dissertation). University of Massachusetts, Amherst.

Schmitz, M. (2008). *The perception of clauses in 6- and 8-month-old German-learning infants. Influence of pause duration and the natural pause hierarchy* (Doctoral dissertation). Universtät Potsdam.

Scott, D. R. (1982). Duration as a cue to the perception of a phrase boundary. *Journal of the Acoustical Society of America*, *71*(4), 996–1007.

Seidl, A. (2007). Infants' use and weighting of prosodic cues in clause segmentation. *Journal of Memory and Language*, *57*(1), 24–48. https://doi.org/10.1016/j.jml.2006.10.004

Seidl, A., & Cristià, A. (2008). Developmental changes in the weighting of prosodic cues. *Developmental Science*, *11*(4), 596–606. https://doi.org/10.1111/j.1467-7687.2008.00704.x

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

Selkirk, E. (1986). On derived domains in sentence phonology. *Phonology*, *3*, 371–405.

Selkirk, E. (2005). Comments on intonational phrasing in English. In S. Frota, M. Vigário, & M. Freitas (Eds.), *Prosodies. With special reference to Iberian languages* (pp. 11–58). Berlin: Mouton de Gruyter.

Selkirk, E. (2011). The Syntax-Phonology Interface. In J. Goldsmith, J. Riggle, & A. C. L. Yu (Eds.), *The Handbook of Phonological Theory* (pp. 435–484). Oxford, UK: Wiley-Blackwell. https://doi.org/10.1002/9781444343069.ch14

Selkirk, E. O., & Tateishi, K. (1988). Constraints on Minor Phrase formation in Japanese. In L. MacLeod, G. Larson, & D. Brentari (Eds.), *Proceedings of CLS 24* (pp. 316–336).

Shukla, M., White, K. S., & Aslin, R. N. (2011). Prosody guides the rapid mapping of auditory word forms onto visual objects in 6-mo-old infants. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(15), 6038–6043. https://doi.org/10.1073/pnas.1017617108

Snedeker, J., & Casserly, E. (2010). Is it all relative? Effects of prosodic boundaries on the comprehension and production of attachment ambiguities. *Language and Cognitive Processes*, *25*(7-9), 1234–1264. https://doi.org/10.1080/01690960903525499

Soderstrom, M., Kemler Nelson, D. G., & Jusczyk, P. W. (2005). Six-month-olds recognize clauses embedded in different passages of fluent speech. *Infant Behavior and Development*, *28*(1), 87–94. https://doi.org/10.1016/j.infbeh.2004.07.001

Soderstrom, M., Seidl, A., Kemler Nelson, D. G., & Jusczyk, P. W. (2003). The prosodic bootstrapping of phrases: Evidence from prelinguistic infants. *Journal of Memory and Language*, *49*(2), 249–267. https://doi.org/10.1016/S0749-596X(03)00024-X

Speer, S. R., & Ito, K. (2009). Prosody in First Language Acquisition - Acquiring Intonation as a Tool to Organize Information in Conversation. *Language and Linguistics Compass*, *3*(1), 90–110. https://doi.org/10.1111/j.1749-818X.2008.00103.x

Steinhauer, K., & Drury, J. E. (2012). On the early left-anterior negativity (ELAN) in syntax studies. *Brain and Language*, *120*(2), 135–162. https://doi.org/10.1016/j.bandl.2011.07.001

Steinhauer, K. (2003). Electrophysiological correlates of prosody and punctuation. *Brain and Language*, *86*(1), 142–164. https://doi.org/10.1016/S0093-934X(02)00542-4

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*(2), 133–138. https://doi.org/10.1016/j.neulet.2010.01.072

Steinhauer, K., Alter, K., & Friederici, A. D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience*, *2*(2), 191–196. https://doi.org/10.1038/5757

Steinhauer, K., & Friederici, A. D. (2001). Prosodic Boundaries, Comma Rules, and Brain Responses: The Closure Positive Shift in ERPs as a Universal Marker for Prosodic Phrasing in Listeners and Readers. *Journal of Psycholinguistic Research*, *30*(3), 267–295.

Streeter, L. A. (1978). Acoustic determinants of phrase boundary perception. *Journal of the Acoustical Society of America*, *64*(4), 1582–1592.

Truckenbrodt, H. (1999). On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry*, *30*, 219–255.

Truckenbrodt, H. (2005). A short report on intonation phrase boundaries in German. *Linguistische Berichte*, *203*, 273–296.

Truckenbrodt, H. (2007a). The syntax-phonology interface. In P. de Lacy (Ed.), *The Cambridge handbook of phonology* (pp. 435–456). Cambridge: Cambridge University Press.

Truckenbrodt, H. (2007b). Upstep on edge tones and on nuclear accents. In C. Gussenhoven & T. Riad (Eds.), *Tones and tunes. Volume 2: Experimental studies in word and sentence prosody* (pp. 349–386). Berlin: Mouton de Gruyter.

Truckenbrodt, H., & Féry, C. (2015). Hierarchical organisation and tonal scaling. *Phonology*, *32*(1), 19–47. https://doi.org/10.1017/S0952675715000032

Vaissière, J. (1983). Language-independent prosodic features. In A. Cutler & D. R. Ladd (Eds.), *Prosody: Models and Measurements* (pp. 53–66). Berlin: Springer.

Wagner, M. (2005). *Prosody and Recursion* (Doctoral Dissertation), Massachusetts Institute of Technology.

Wagner, M. (2010). Prosody and recursion in coordinate structures and beyond. *Natural Language & Linguistic Theory*, *28*(1), 183–237. https://doi.org/10.1007/s11049-009-9086-0

Wagner, M., & Crivellaro, S. (2010). Relative Prosodic Boundary Strength and Prior Bias in Disambiguation. In M. Hasegawa-Johnson (Ed.), *Proceedings of Speech Prosody 5.* Chicago, IL.

Wagner, M., & Watson, D. G. (2010). Experimental and theoretical advances in prosody: A review. *Language and Cognitive Processes*, *25*(7-9), 905–945. https://doi.org/10.1080/01690961003589492

Watson, D., & Gibson, E. (2004). The relationship between intonational phrasing and syntactic structure in language production. *Language and Cognitive Processes*, *19*(6), 713–755. https://doi.org/10.1080/01690960444000070

Wellmann, C., Holzgrefe, J., Truckenbrodt, H., Wartenburger, I., & Höhle, B. (2012). How each prosodic boundary cue matters: evidence from German infants. *Frontiers in Psychology*, *3*, 580. https://doi.org/10.3389/fpsyg.2012.00580

Wellmann, C., Holzgrefe-Lang, J., Truckenbrodt, H., Wartenburger, I., & Höhle, B. (submitted). Prosodic boundary cue perception in German infants: Developmental change from six to eight months.

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America*, *91*(3), 1707–1717.

Wolff, S., Schlesewsky, M., Hirotani, M., & Bornkessel-Schlesewsky, I. (2008). The neural mechanisms of word order processing revisited: electrophysiological evidence from Japanese. *Brain and Language*, *107*(2), 133–157. https://doi.org/10.1016/j.bandl.2008.06.003

Yang, X., Shen, X., Li, W., & Yang, Y. (2014). How listeners weight acoustic cues to intonational phrase boundaries. *PloS ONE*, *9*(7), e102166. https://doi.org/10.1371/journal.pone.0102166

Yu, A. (2010). Tonal effects on perceived vowel duration. In C. Fougeron, B. Kuehnert, M. Imperio, & N. Vallee (Eds.), *Papers in Laboratory Phonology* (Vol. 10, pp. 151–168). Berlin, New York: de Gruyter.

Zhang, X. (2012). *A comparison of cue-weighting in the perception of prosodic phrase boundaries in English and Chinese* (Doctoral dissertation). University of Michigan.

APPENDIX

Table A.1 | Complete list of experimental stimuli with acoustic key characteristics.

Cond.	Name sequence (critical name in italics, critical final vowel in bold)	f0 rise on critical name in Hz	f0 maximum on final vowel in Hz	final vowel duration in ms	total duration in ms
NO	1: Manni und *Len**i*** und Nina	90.72	275.42	108	1796
	2: Moni und *Lill**i*** und Manu	77.82	270.82	95	1779
	3: Lilli und *Mann**i*** und Nina	87.82	273.22	125	1803
	4: Leni und *Mimm**i*** und Manu	84.16	286.86	101	1829
	5: Nelli und *Mon**i*** und Lola	65.09	251.89	103	1780
	6: Mimmi und *Nell**i*** und Lola	89.20	274.10	95	1842
	Mean	**82.47**	**272.05**	**105**	**1805**
PI	1: Manni und *Len**i*** und Nina	199.29	379.39	108	1796
	2: Moni und *Lill**i*** und Manu	209.90	389.60	95	1779
	3: Lilli und *Mann**i*** und Nina	208.84	389.24	125	1803
	4: Leni und *Mimm**i*** und Manu	202.99	386.39	101	1829
	5: Nelli und *Mon**i*** und Lola	207.53	387.73	103	1780
	6: Mimmi und *Nell**i*** und Lola	206.13	386.13	95	1842
	Mean	**205.78**	**386.41**	**105**	**1805**
PIDUR	1: Manni und *Len**i*** und Nina	210.28	391.08	191	1883
	2: Moni und *Lill**i*** und Manu	211.54	391.64	186	1856
	3: Lilli und *Mann**i*** und Nina	212.87	394.07	207	1904
	4: Leni und *Mimm**i*** und Manu	209.38	391.98	175	1911
	5: Nelli und *Mon**i*** und Lola	210.68	392.28	185	1863
	6: Mimmi und *Nell**i*** und Lola	211.44	391.64	178	1919
	Mean	**211.03**	**392.11**	**187**	**1889**
PI2DUR	1: Manni und *Len**i*** und Nina	211.29	391.99	191	1883
	2: Moni und *Lill**i*** und Manu	211.41	391.51	186	1856
	3: Lilli und *Mann**i*** und Nina	213.97	393.87	207	1904
	4: Leni und *Mimm**i*** und Manu	212.54	392.44	175	1911

	5: Nelli und *Mon**i*** und Lola	211.77	392.07	185	1863
	6: Mimmi und *Nell**i*** und Lola	211.52	391.72	178	1919
	Mean	**212.08**	**392.27**	**187**	**1889**
DUR	1: Manni und *Len**i*** und Nina	89.52	275.52	191	1883
	2: Moni und *Lill**i*** und Manu	79.42	271.82	186	1856
	3: Lilli und *Mann**i*** und Nina	86.95	272.35	207	1904
	4: Leni und *Mimm**i*** und Manu	84.22	287.92	175	1911
	5: Nelli und *Mon**i*** und Lola	65.85	253.55	185	1863
	6: Mimmi und *Nell**i*** und Lola	92.31	275.21	178	1919
	Mean	**83.04**	**272.73**	**187**	**1889**
PI2	1: Manni und *Len**i*** und Nina	199.35	379.65	108	1796
	2: Moni und *Lill**i*** und Manu	210.17	390.07	95	1779
	3: Lilli und *Mann**i*** und Nina	209.18	389.38	125	1803
	4: Leni und *Mimm**i*** und Manu	203.18	386.28	101	1829
	5: Nelli und *Mon**i*** und Lola	208.13	388.13	103	1780
	6: Mimmi und *Nell**i*** und Lola	206.89	386.49	95	1842
	Mean	**207.07**	**386.67**	**105**	**1805**

Note: Slight deviations in the f0 values across conditions with pitch manipulations are due to PSOLA resynthesis and/ or measuring inaccuracies induced by PRAAT.

Erklärung

Hiermit erkläre ich, Julia Holzgrefe-Lang, dass ich die vorliegende Dissertation selbstständig und ohne Hilfe Dritter verfasst habe. Bei der Abfassung der Arbeit wurden alle Regelungen guter wissenschaftlicher Standards eingehalten. Ich erkläre ferner, dass ich diese Arbeit ohne Benutzung anderer als der angegeben Quellen und Hilfsmittel verfasst habe.

Leipzig, den 3. Februar 2017 ______________________________