

Prosodic processing at the sentence level in infants

Ann Pannekamp^{a,b}, Christiane Weber^a and Angela D. Friederici^a

^aMax Planck Institute for Human Cognitive and Brain Sciences, Leipzig and ^bHumboldt University, Berlin, Germany

Correspondence and requests for reprints to Angela D. Friederici, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstrasse 1a, 04103 Leipzig Germany
Tel: + 49 341 9940 112; fax: + 49 341 9940 113; e-mail: angelafr@cbs.mpg.de

Sponsorship: This research was supported by the Deutsche Forschungsgemeinschaft (German Research Foundation, DFG) (FR-519/18-1) and by the Schram Foundation (T278/10824/2001, Genetic Bases of Specific Language Impairment).

Received 22 December 2005; accepted 3 February 2006

We investigate whether 8-month-old infants can detect prosodic cues relevant in sentence structuring. We recorded event-related potentials to examine online responses to the processing of prosodic boundaries. Prior studies in adults have validated the closure positive shift as reflecting prosodic boundary perception during speech processing. The current study shows that in the event-related potentials of 8-month-olds, a closure positive shift is elicited in relation to the prosodic boundaries in speech,

suggesting that these infants are able to structure speech input into prosodic units on a neurophysiological basis similar to that seen in adults. A delay in latency of the infant closure positive shift, however, suggests that children's exploitation of prosodic boundaries for the segmentation of the speech stream is still developing. *NeuroReport* 17:675–678 © 2006 Lippincott Williams & Wilkins.

Keywords: closure positive shift, event-related potentials, infants, language acquisition, prosody

Introduction

Prosody is an inherent aspect of spoken language and an important factor in human verbal interaction. In particular, prosody can be relevant for differentiating the meanings of words (e.g. *record* vs. *recorder*) and sentences (e.g. questions vs. statements). Moreover, prosody contributes substantially to the segmentation of the spoken speech stream into single utterances and into more fine-grained prosodic phrases. These prosodic phrases are not only tonal units, but they also correspond to major syntactic phrases. This structural correspondence applies, *inter alia*, to the prosodic phrase type of intonational phrases (henceforth IPhs) [1]. The boundaries of IPhs are defined by several characteristics. First, the last syllable of an IPh contains a distinctive boundary tone, which can be either high or low. Second, this last syllable is longer than syllables at within-phrase positions. Last, a pause can follow an IPh. Many behavioral studies show that IPh boundaries are used by adult listeners to interpret and disambiguate spoken utterances [2–4].

Recently, several neurophysiological studies have explored the influence of prosodic information on lexical processing during speech perception [5,6]. For syntactically relevant IPhs, a particular event-related potential (ERP) response called the closure positive shift (CPS) has been identified [7]. This component has a centro-parietal scalp distribution and peaks at a latency of approximately 500 ms after the IPh boundary. In the present study, the CPS is interpreted as reflecting the closure of an IPh. The CPS component has also been observed in relation to musical phrasing [8] and the perception of prosodic boundaries in speech-like materials lacking segmental content [9].

The prosodic aspects of language are ascribed a substantial role in first-language acquisition as the segmentation of the speech stream is a necessary prerequisite for the identification of structural units in the input. Several studies to date have attested that even newborns are able to distinguish the speech melody of their mother tongue from a language with diverging sentence prosodies [10–13]. On the basis of these early language abilities, children learn to derive syntactic ordering principles quite automatically. This process has been termed 'prosodic bootstrapping' [14]. Behavioral experiments indicate that, at around 9 months of age, infants are capable of recognizing major syntactic phrases in the acoustic input on the basis of prosodic information [15,16]. The current study explores whether infants have a neurophysiological basis similar to adults for online detection of prosodic boundaries that cue syntactic ordering principles. If adults and infants are similar, then a certain amount of developmental continuity of the language processes [17] can be assumed. If infants show a very different pattern from adults, then the underlying processes at an early developmental stage can also be assumed to differ [18].

Subjects and methods

Participants

The experiment was part of the longitudinal German Language Development Study conducted at the Children's Hospital Lindenhof in Berlin. Families participated in this study according to institutional informed consent procedures. Before the experiment, the infants' hearing was tested

by an otoacoustic emission screening. All infants had normal hearing and no health, developmental, or neurological problems. Thirty-four infants (11 female) took part in the present experiment. They were between 31 and 35 weeks old (mean 32.7 weeks).

Stimulus materials

The sentence materials consisted of two conditions (see Table 1). Sentences consisted of a critical part that differed in conditions A and B, and a noncritical part that served as a sentence continuation that is identical in both conditions and which ensured that there was no abrupt acoustic offset at the end of the critical sentence part. The two conditions diverge with respect to the valence of the second verb. In condition A, the second verb *zu schlafen* (to sleep) is intransitive. Thus, the noun phrase *Mama* is the object of the first verb *verspricht* (promise). In turn, the only IPh exists after the critical verb *zu schlafen* in *Kevin verspricht Mama zu schlafen* (Kevin promises Mom to sleep). In contrast, the second verb *zu küssen* (to kiss) in condition B is transitive and takes the noun *Mama* as its object. As a result, one IPh occurs at the second verb constituting a first syntactic phrase *Kevin verspricht* (Kevin promises), and a second phrase *Mama zu küssen* (to kiss Mom). Thus, the two sentence types differ in their syntactic structure as well as in their prosodic realization.

Each condition consisted of 48 sentences. All utterances were produced by a trained female speaker using infant-directed speech. The recordings took place in a soundproof chamber and were digitized (44.1 kHz/16 bit sampling rate/mono). All sentences were normalized and adjusted in amplitude.

The analysis of the tonal and durational properties of the sentence materials shows clear differences between the two conditions. Only in the transitive condition (condition B) does the fundamental frequency pattern exhibit a high tone at the verb *verspricht*. After the end of the critical sentence part, both conditions show an IPh boundary, indicated mainly by a high boundary tone. Another difference between conditions is the duration of the first two words of the sentence, which is significantly longer for condition B as shown by a two-tailed t-test [$t(94) = -6.00, P < 0.01$]. Furthermore, condition B exhibits a significantly longer pause after the first two words in comparison to condition A [$t(94) = -21.86; P < 0.01$]. In summary, Condition A displays no IPh boundary in the critical part of the sentence, whereas condition B exhibits one IPh boundary in the critical part (at 950 ms postsentence onset). The mean length of the entire

sentence for condition A was 3800 ms; for condition B, it was 4030 ms.

Procedure and event-related potential recording

The sentence stimuli were presented acoustically while the infants listened passively, seated on their mother's or father's lap. The two conditions were presented in a pseudo-randomized order and in blocks of 12 sentences. The order of the blocks was counterbalanced across children.

The electroencephalogram was continuously recorded from 23 Ag/AgCl electrodes attached to an elastic electrode cap (Easy Cap, Falk Minow Services, Herrsching Breitenbrunn, Germany). Electrodes were placed at F7, F3, FZ, F4, F8, FT7, FC3, FC4, FT8, T7, C3, CZ, C4, T8, CP5, CP6, P7, P3, PZ, P4, P8, O1, and O2. The electro-oculogram was recorded from electrodes at the outer canthus of each eye and from sites above and below the right eye. The system was referenced online to the left mastoid and re-referenced offline against linked mastoids. Following the recordings, a 0.3 Hz high-pass filter was applied to the data sets.

Data analysis

Electroencephalogram epochs containing eye-blinks or movement artifacts were rejected and did not enter the ERP averages. After the rejection procedure, at least 10 trials per condition and infant entered the following analyses. Averages were computed with a duration of 3000 ms covering the critical part of the stimuli and with a 200 ms prestimulus baseline. As infants' attention decreased towards the end of the sentence, only the critical part of the sentence was analyzed. First, averages per condition per child were calculated, and then condition means were calculated across all children.

For the statistical analysis, analyses of variance (ANOVAs) were computed for the midline and lateral electrodes separately for each experimental condition. At the midline electrodes (FZ, CZ, and PZ), a two-way ANOVA was performed with the factors condition (A vs. B) and region (frontal vs. central vs. parietal). For the lateral electrodes, six regions of interest were defined by crossing the two factors region (frontal vs. central vs. parietal) and hemisphere (left vs. right). A three-way ANOVA design with the factors condition (A vs. B), region (frontal vs. central vs. parietal), and hemisphere (left vs. right) was then conducted. All ANOVAs were computed in time windows of 500 ms.

Results

In Fig. 1, ERPs for both conditions are displayed from the sentence onset until the second verb, which ends the only phrase in condition A and the second phrase in condition B in which a first IPh occurs after the first verb. It becomes evident from Fig. 1 that this IPh in condition B (the transitive condition) elicits a positive shift, while in condition A (the intransitive condition) there is no IPh present after the first verb. The positive ERP waveform for condition B starts at about 1700 ms and reaches a maximum between 2000 and 2500 ms postsentence onset. The statistical analysis reveals a main effect of condition in the time window 2000–2500 ms at the midline [$F(1,33) = 12.35, P < 0.05$] as well as at the lateral electrodes [$F(1,33) = 4.44, P < 0.05$]. No significant effects were observed later in the sentences.

Table 1 Examples of the materials (with literal translations)

	Critical part	Noncritical part
Condition A		
No IPh in critical part (intransitive)	(Kevin verspricht Mama zu schlafen) (Kevin promises Mom to sleep)	(und ganz lange lieb zu sein.) (and to be a good boy for a while.)
Condition B		
One IPh in critical part (transitive)	(Kevin verspricht,) IPh (Mama zu küssen) (Kevin promises,) IPh (to kiss Mom)	(und ganz lange lieb zu sein.) (and to be a good boy for a while.)

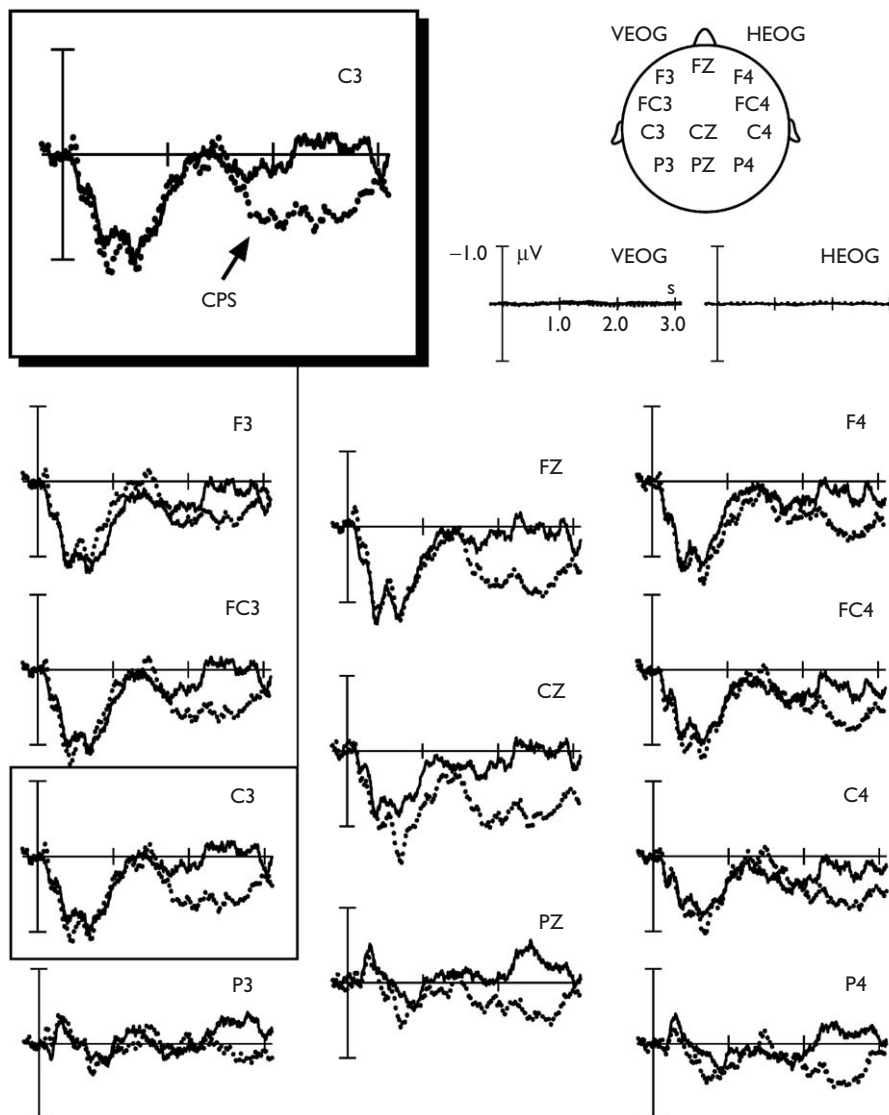


Fig. 1 Event-related potentials (ERPs) elicited from 8-month-olds for sentence parts with different prosodic contours time-locked to the onset of the sentence are displayed. Condition B with one intonational phrase (IPh) boundary in its critical part (transitive verb) is contrasted with condition A with no IPh boundary in its critical part (intransitive verb). The perception of the IPh boundary in condition B leads to a positive-going ERP at around 2000 ms, while in condition A the boundary and ERP response are not present.

Discussion

The ERP data from 8-month-old infants showed clear differences between the two sentence types with differing syntactic and prosodic realizations, respectively. While the perception of the IPh boundary evoked a positive shift, sentences without a major prosodic boundary in this particular position did not. As a similar positive shift has already been observed in adults at IPhs, the present effect was interpreted as reflecting the processing of IPh boundaries in infants. The latency of the infants' response, however, is delayed as compared with prior findings in adults. The delayed onset of ERP components is not unexpected, given that delayed onsets of ERP components in early infancy have also been observed for the processing of semantic [19] and syntactic information [20].

The present finding supports the view that prosodic processes crucial for speech segmentation are established very early in development. Using behavioral methods, it has

been shown that 9-month-old infants react to the acoustic cues correlated to major phrase boundaries [16], and that 10-month-olds can use phonological phrase boundaries to segment speech [21]. Here, we demonstrate that the neural basis for the perception of prosodic phrase boundaries is established as early as 8 months. The similarity of the observed ERP pattern between 8-month-old infants and adults suggests that the underlying processes are the same in principle, though somewhat slower in early infancy than in adulthood. Therefore, the present results may be taken to support the view of developmental continuity in language processing.

Conclusion

The present ERP data suggest that infants at the age of 8 months are able to use major prosodic phrase boundaries for the segmentation of connected speech, and that this ability

is established during early infancy with a neurophysiological basis that is similar to that of adults. Although there is a delay in the onset of the prosody-related ERP component in young infants suggesting that speech segmentation is emerging, the general similarity of this component in infants and adults indicates developmental continuity of the mechanisms underlying language processing.

Acknowledgements

We wish to thank Christina Rügen and Jödis Haselow for recording the ERP data, and all the families who participated in the study. We also thank Ulrike Toepel for her support in the prosodic analyses.

References

- Selkirk E. *Phonology and syntax: the relation between sound and structure*. Cambridge, MA: MIT Press; 1984.
- Kennedy A, Murray WS, Jennings F, Reid C. Parsing complements: comments on the generality of the principle of minimal attachment. *Lang Cogn Process* 1984; **4**:51–76.
- Pynte J, Prieur B. Prosodic breaks and attachment decisions in sentence parsing. *Lang Cogn Process* 1996; **11**:165–192.
- Cutler A, Dahan D, van Donselaar W. Prosody in the comprehension of spoken language: a literature review. *Lang Speech* 1997; **40**:141–201.
- Holcomb PJ, Neville HJ. Natural speech processing: an analysis using event-related potentials. *Psychobiology* 1991; **19**:286–300.
- Hayashi R, Imaizumi S, Mori K, Niimi S, Ueno S, Kiritani S. Elicitation of N400m in sentence comprehension due to lexical prosody incongruity. *NeuroReport* 2001; **12**:1753–1756.
- Steinhauer K, Alter K, Friederici AD. Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nat Neurosci* 1999; **2**:191–196.
- Knösche TR, Neuhaus C, Hauelsen J, Alter K, Friederici AD, Witte OW. The perception of phrase structure in music. *Hum Brain Mapp* 2005; **24**:259–273.
- Pannekamp A, Toepel U, Alter K, Hahne A, Friederici AD. Prosody-driven sentence processing. *J Cogn Neurosci* 2005; **17**:407–421.
- Mehler J, Jusczyk P, Lambertz G, Halsted M, Bertoncini J, Amiel-Tison C. A precursor of language acquisition in young infants. *Cognition* 1988; **29**:143–178.
- Moon C, Cooper RP, Fifer WP. Two-days-olds prefer their native language. *Infant Behav Dev* 1993; **16**:495–500.
- Nazzi T, Floccia C, Bertoncini J. Discrimination of pitch contours by neonates. *Infant Behav Dev* 1998; **21**:779–784.
- Ramus F, Hauser MD, Miller C, Morris D, Mehler J. Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science* 2000; **288**:349–351.
- Gleitman LR and Wanner E. Language acquisition: the state of the state of the art. In: Wanner E, Gleitman L, editors. *Language acquisition: the state of the art*. New York: Cambridge University Press; 1982. pp. 3–48.
- Hirsh-Pasek K, Kemler-Nelson DG, Jusczyk P, Wright-Cassidy K, Druss B, Kennedy L. Clauses are perceptual units for young infants. *Cognition* 1987; **26**:269–286.
- Jusczyk P, Hirsh-Pasek K, Kemler-Nelson DG, Kennedy L, Woodward A, Piwoz J. Perception of acoustic correlates to major phrasal units by young infants. *Cogn Psychol* 1992; **24**:252–293.
- Pinker S. *Language learnability and language development*. Cambridge, MA: Harvard University Press; 1984. pp. 37–62.
- Felix S. Maturational aspects of universal grammar. In: Davis A, Crippen C, Howatt A, editors. *Interlanguage*. Edinburgh: Edinburgh University Press; 1994. pp. 133–161.
- Friedrich M, Friederici AD. Semantic sentence processing reflected in the event-related potentials of one- and two-year old children. *NeuroReport* 2005; **16**:1801–1804.
- Oberecker R, Friedrich M, Friederici AD. Neural correlates of syntactic processing in two-year-olds. *J Cogn Neurosci* 2005; **17**:1667–1678.
- Grout A, Christophe A, Morgan JL. Phonological phrase boundaries constrain lexical access: II. Infant data. *J Mem Lang* 2004; **51**:548–567.