Don't blame it (all) on the pause: Further ERP evidence for a prosody-induced garden-path in running speech

Karsten Steinhauer, Kai Alter and Angela D. Friederici

Max-Planck-Institute of Cognitive Neuroscience, Leipzig, Germany

steinhau@cns.mpg.de

ABSTRACT

This paper investigates the prosodic relevance of a pause which, along with other prosodic parameters, served to indicate an Intonational Phrase (IPh) boundary. Event-related brain potentials (ERPs) were recorded while subjects listened to both intact and altered German Early and Late Closure (EC/LC) sentences. The EC sentences were prosodically highly accepted and well comprehended even when the original pause at the boundary position was removed. Furthermore, a reversed garden-path (initial EC preference in LC sentences) was successfully induced by a false IPh boundary irrespective of whether the pause was present or not. The ERP patterns disclosed the on-line processing of simple and garden-path sentences in more detail. The data clearly demonstrate that in the presence of other prosodic parameters pause insertion is a completely dispensable cue for boundary marking. The ERP technique proved to be superior to behavioral on-line measures as data collection does not interrupt speech presentation.

1. The Separation of Prosodic Parameters

Speech, in contrast to written text, provides prosodic cues in order to express both linguistic (e.g. syntactic) and non-linguistic (e.g. affective) information. However, in order to realize a certain prosodic effect (e.g. accentuation, prosodic phrasing, etc.), speakers can use a variety of prosodic parameters such as pause insertion, constituent lengthening, and pitch or loudness variations (Cutler, Dahan & Donselaar, 1997; see also Alter, Steinhauer & Friederici, this issue). The hearer, on the other hand, has to decode and to integrate these different parameters. A central question among the 'HOWs' of prosodic language processing thus concerns the relative contribution of each single parameter. A related issue deals with the rules according to which the parameters can or must be combined in order to achieve a certain effect. A crucial prerequisite for addressing these questions empirically is the separation and systematic variation of single parameters. This is a non-trivial task as the different parameters are usually not completely independent of each other. A relatively independent prosodic parameter is the duration of pauses. Of course, pauses cannot be simply inserted at any position as this would destroy the internal structure of intonational units and/or disrupt co-articulation. However, wherever a pause was originally produced by a speaker, its duration can in principle be manipulated without affecting other parameters. Such manipulations were successfully carried out in the materials of the present study in order to examine the relative contribution of a pause to the prosodic realization of IPh boundaries.

2. When Prosody meets Syntax

The pauses referred to had been observed in the Early Closure condition of a previous auditory study investigating the prosodic processing of German Late and Early Closure (LC/EC) ambiguities as illustrated in (A) and (B) and described in detail below. In (A), the second verb *'arbeiten'/'to work'* is intransitive and NP2 *'Anna'* is the indirect object of the preceding verb_1 *'verspricht'/'promises'* (LC). In (B), by contrast, *'Anna'* is demanded as direct object by the subsequent transitive verb_2 *'entlasten'/'to support'* (EC). Note that both conditions are structurally ambiguous up to the verb_2 which is intransitive in (A) and transitive in (B).

Late Closure (*'Anna'* as indirect object of verb_1):

(A) **[IPh1** *Peter verspricht Anna zu arbeiten*]
Peter promises Anna to work

[IPh2 *und das Büro zu putzen.*]
and to clean the office

Early Closure (*'Anna'* as direct object of verb2):

(B) **[IPh1** *Peter verspricht #*] **[IPh2** *Anna zu entlasten*]
Peter promises # to support Anna

[IPh3 *und das Büro zu putzen.*]
and to clean the office

Comprehensive acoustic analyses of each of the 48 LC/EC sentence pairs had disclosed that the speaker had expressed the structural differences between (A) and (B) by very early prosodic differences in the speech signals. Similar to findings reported by Warren.....ET AL.. (1995a,b) for English EC/LC ambiguities, these prosodic differences occurred before the sentences were structurally disambiguated by the argument structure of verb_2. As a consequence, the garden-path effect in EC sentences predicted by the Late Closure Principle proposed by Frazier (e.g. Frazier & Rayner, 1982) was found only for visual presentation in a reading task. When presented auditorily, however, the early availability of disambiguating prosodic information secured the immediate appropriate analysis in both the LC and the EC condition and thus prevented the listeners from the garden-path.

The pause insertion under consideration (marked by a '#' in example (B)) was one of the most salient prosodic differences between the two conditions. They were obviously produced by the speaker in order to mark the additional IPh boundary between the first verb *'verspricht'* and NP2 *'Anna'* in EC

sentences. This pattern is in complete agreement with the predictions of certain theories of syntax-prosody-mapping (Reyelt et al., 1996). In other words, the early *syntactic* closure (EC) was prosodically realized by the early closure of the first Intonational Phrase.

For the hearer, the detectability of this additional IPh boundary served as a highly important criterion for parsing decisions. That is, its presence determined an initial EC analysis instead of the otherwise preferred LC analysis. The dominance of prosodic information in guiding initial parsing decisions was most convincingly demonstrated when we introduced the additional IPh boundary of (B) via cross-splicing into the LC condition (A): Although the resulting condition (C) consisted of the same lexical elements as the normally easy-to-process condition (A), the conflict between the early IPh boundary (signaling EC) and the demands of the intransitive verb (requiring LC) induced a very severe garden-path (Steinhauer, Alter & Friederici, 1998).

False Early Closure (intransitive verb2 cannot take an object):

(C) * |**IPh1** *Peter verspricht #*| |**IPh2** *Anna zu arbeiten*|
* Peter promises # to work Anna

|**IPh3** *und das Büro zu putzen.*|
and to clean the office

Interestingly, this prosody-induced garden-path was the *reverse* version of usual LC/EC garden-paths as it required an *initial EC analysis* to be revized towards an LC interpretation, and not vice versa. Such sentences were initially perceived as *'Peter promises to work Anna ...'* which is certainly not grammatical. The on-line effects of both the prosodic phrasing and the garden-path effect (including structural reanalyses) were clearly reflected by event-related brain potentials (ERPs).

3. The Brain at Work: Event-related Potentials and Language Processing

Most psycholinguistic research is based on behavioral studies measuring error rates and reaction times. However, due to its implicit on-line characteristics and its high time resolution the employment of ERP measures has joined the list of on-line methods as an additional approach to study language processing. ERPs are a transient change of voltage, reflecting a systematic brain activity which is triggered by a physical event. If this event is a word presented either in a semantically appropriate context or in an inappropriate context (e.g. *'He spread the warm bread with* ***socks****'*), the ERP differences between the two conditions reflect the brain's activity while processing a semantic violation. This experiment was conducted by Kutas and Hillyard (1980) who found a negative brain potential around 400 ms after onset of the anomalous word, the N400 component. The N400 has been replicated in hundreds of studies and reflects difficulties in lexical/semantic integration. The costs of processing syntactic difficulties, on the other hand, generally elicit a late positivity around 600 ms (P600) rather than an N400 (e.g. Osterhout & Holcomb, 1992). Thus, in contrast to response times, ERP patterns can more easily distinguish between different levels of linguistic processing.

The verb argument structure violations in the reverse garden-path condition (C) imply both an initial problem of lexical integration and additional syntactic cost while reanalyzing the structure from EC to LC. As predicted, the incompatible verb_2 of condition (C) elicited an N400 followed by a P600. As in comparable auditory studies presenting natural speech, both ERP compononts had a later peak latency as compared to visual studies such as those cited above. A completely new finding of our previous study was that IPh boundaries were also reflected in the ERPs by a positive going waveform that we refer to as Closure Positive Shift (CPS) (Steinhauer et al., 1998).

4. The Present Study

After the previous studies described above had proved that prosody drove the parsing decisions we could prepare to examine the relative contribution of single prosodic parameters. The most salient prosodic differences between the speech signals in (A) and (B) were as follows:

- First, the initial sentence segment (i.e. *'Peter verspricht'*) was considerably lengthened in (B).
- Second, a pause of some 150 ms was inserted only in (B) at the additional IPh boundary (i.e. before *'Anna'* as marked by a '#').
- Third, whereas the main accent in (A) was aligned to the second verb *'arbeiten'*, it occurred on the NP*'Anna'* in (B). The respective accent positions were realized by both pitch and loudness peaks.

As outlined above, the most independent one of these parameters was the pause duration. The present study therefore focuses on potential parsing differences due to the presence as opposed to the absence of this pause. To this aim, we manipulated the speech signals of the former conditions (B) and (C) in that we removed the pause between the first verb (e.g. *'verspricht'/'promises'*) and the second NP (e.g. *'Anna'*). The pause removal was performed with particular care in order not to alter the signals of adjacent words. For example, as stop consonants such as /p/ are characterized by a preceding silent phase of some 50 ms, the corresponding part of the pause was preserved whenever the NP2 began with a stop consonant.

The two resulting conditions derived from (B) and (C) will hereafter be referred to as (B") and (C"), respectively. The study comprised 4 experimental conditions, namely **(A)**, **(B)**, **(B")**, and **(C")** with 48 sentences each. Note that (C") consisted of the same lexical elements as (A) including the intransitive verb requiring LC; and (B") consisted of the same elements as (B) with the transitive verb requiring EC. In contrast to (A), the initial fragments in conditions (B), (B"), and (C") were all derived from the EC condition and thus contained the same pitch and loudness patterns. However, only the original condition (B) still contained a pause preceding NP2 (e.g. *'Anna'*), whereas (A), (B") and (C") did not. The relevant parts of the respective 4 conditions are enumerated below, where *italics* indicate the origin from the LC condition (A) and **bold** characters indicate the origin from EC condition (B). The '#'

again signals the presence of a pause, and SMALL CAPITALS indicate the word carrying the respective main accent.

(A) *Peter verspricht Anna zu ARBEITEN*

(B) **Peter verspricht # ANNA zu entlasten**

(B") **Peter verspricht ANNA zu entlasten**

(C") **Peter verspricht ANNA** *zu arbeiten*

The general rationale of this design is as follows. The detectablity of the additional IPh boundary between *'verspricht'* and *'Anna'* is held responsable for the change from initial LC to initial EC parsing. If the pause insertion as such contributes considerably to the boundary marking, then its absence should reduce this change. That is, (B") should be more difficult to parse than (B) as it relies on the detection of the boundary. And (C") should induce a less severe garden-path than the former (C) condition as its intransitive verb_2 *'arbeiten'* requires a final LC analysis.

4.1. Materials and Methods

Subjects. 16 students participated in the experiment. All subjects were right-handed German native speakers with no neurological disorders. They were not informed in advance about the aims of the study. None of them had experience with related studies.

Materials. The 4 × 48 experimental sentences were pseudo-randomly intermixed with 144 filler sentences (produced by the same female native speaker of Standard German) and distributed across 8 experimental blocks. The blocks were presented in two sessions with 4 blocks each. For 20 % of the sentences, comprehension questions were produced by a male speaker.

Procedure. After electrode application (see below) participants were seated in a sound-proof and electro-magnetically shielded dimly lit chamber. They sat 80 cm away from a 17" monitor and two loudspeakers. After performing in a practice block of 10 trials they were presented with the four experimental blocks which were separated by pauses. Each trial began with a fixation cross in the center of the monitor in order to minimize eye movements which can cause artifacts in the EEG. After 2000 ms, the sentence was presented auditorily, while the cross remained on the screen. Subjects were instructed to listen carefully to the sentences. Immediately after sentence presentation, the question 'correct?' was presented visually on the screen and the participants had to judge the prosodic acceptability of the sentence by pressing a YES or a NO button. In 20 % of the trials, an additional comprehension question was presented auditorily and had to be answered as accurately and fast as possible. Thereafter, an exclamation mark on the screen indicated a 2000 ms interval in which subjects were encouraged to blink their eyes. This instruction reduced eye blink artifacts during sentence presentation. Trial presentation and performance data collection was controlled by an IBM compatible Pentium PC using ERTS software.

EEG recording. The EEG was recorded from 17 cap-mounted tin electrodes with a sampling rate of 250 Hz/12 bits and amplified by a Neuroscan DC amplifier with 40 Hz low-pass filter. The left mastoid electrode served as the reference.

ERP analyses. ERPs for each participant, condition, and electrode were computed for epochs time-locked either to the sentence onset or to the onset of the critical verb_2 with preceding baseline intervals of 200 ms. Trials containing eye blink or movement artifacts were rejected. Averages were first computed for each single subject. These averages then entered the grand averages. ERP components were quantified as amplitude means of specified time windows. (For details see the Results section.) Statistical analyses for both behavioral and ERP data were performed by ANOVAs. Where appropriate, Huyhn & Feldt df-corrections and a modified Bonferroni p-value correction to protect agains progressive Type-I errors were applied. All effects were significant with $p<.05$ or better.

4.2. Hypotheses

Given the findings of the previous study, the hypotheses were straight forward. As general hypotheses we expected a replication for conditions (A) and (B):

- High acceptabilities and low error rates in both conditions.
- Closure Positive Shifts in the ERP at IPh boundaries, i.e. one CPS in (A) and two CPS's in (B)

Concerning the pause manipulation, two possible cases had to be distinguished from which the following hypotheses could be derived:

1.) **If** the pause insertion in the original EC condition (B) was crucial or at least important for IPh boundary marking and to induce EC parsing, **then** ...

1.a. (B") should be considerably less acceptable than (B), and processing difficulties when encountering verb_2 in (B") should elicit an N400/P600 pattern in the ERP.

1.b. If the boundary marking was completely due to the pause (C") should be as well accecpted as (A) and display the same ERP patterns. If the pause was only partly contributing to the boundary marking, (C") should be more readily accepted than (C) of the previous study and elicit a smaller N400/P600 pattern in the ERP than (C)

2.) **If**, however, the pause insertion in (B) was *not* relevant for IPh boundary marking and to induce EC parsing, **then** ...

2.a. (B") should be equally acceptable as (B) and the verb_2 should elicit the same ERP pattern as in B).

2.b. (C") should be just as prosodically acceptable as (C) of the previous study and elicit the same N400/P600 pattern in the ERP as (C).

5. Results

The results were almost as straight forward as the hypotheses. As **Figure 1** illustrates, the prosodic judgement data revealed a clear pattern. Both condition (A) and (B) were highly accepted in 82.6 and 74. 1 % of the trials and did not differ from each other, thus confirming the first general hypothesis. (B") was equally well accepted (75.8 %) and did not differ from (B), confirming hypothesis 2a and rejecting 1a. (C") was accepted in only 14.5% of the trials. This rate was significantly smaller than that in (A) but still higher than that of the original (C) condition in the previous study. This pattern at least rejects the strongest version of 1b.

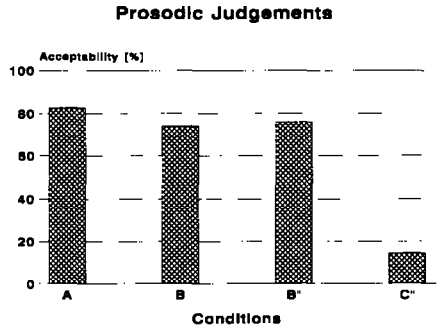

Figure 1: Prosodic Judgements in the four conditions, given as acceptability rates.

Second, an early positive going shift in the ERPs at the initial IPh (i.e. the first CPS) was found in condition (B), (B") and (C"), but not in (A). The late CPS was observed in all conditions. This ERP pattern confirms the second general hypothesis. Furthermore, it shows that even without a pause the IPh was processed on-line. Note, that due to the pause removal, the first CPS by 1 second in conditions (B") and (C") had a shorter latency than that in (B).

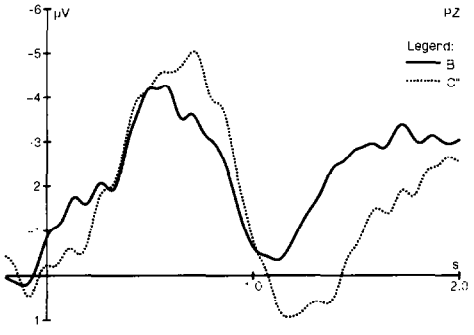

Figure 2: Grand average ERPs time-locked to the onset of the critical verb_2 at the PZ electrode. The potentials of conditions (B) and (C") are superimposed.

Third, the ERPs elicited by the second verb do not differ between conditions (B) and (B"), indicating similar on-line processing. In contrast, verb_2 in condition (C") displays considerably enhanced amplitudes of both N400 and P600 (**Figure 2**). This pattern is identical to that previously found for the original (C) condition. This finding confirms hypothesis 2b while rejecting 1b.

6. Discussion

The data clearly demonstrate that in the presence of other prosodic parameters, a the original pause insertion was completely dispensable for the hearers' detection of the additional first IPh boundary. Even when the pause was removed, EC instead of LC parsing was initially determined. This parsing decision led to easy integration of verb_2 in (B") and to a reverse garden-path in (C"). The strength of this reverse garden-path seems to indicate that a reanalysis involving an IPh boundary deletion is extremely difficult. Its severeness may also have to do with the revision of the initally assumed dominance relation between verb_2 and NP2 (cf. Gorrell, 1995).

In contrast to purely behavioral measures, the CPS in the ERPs provides an on-line indication of intonational phrasing which otherwise can only be indirectly inferred. The N400/P600 pattern characterizes the processes underlying the garden-path and their respective time course more specifically than behavioral data. An additional advantage of ERPs was that the sentences could be presented as a whole. Behavioral on-line studies employing cross-modal naming tasks (e.g. Warren et al., 1995), present sentence fragments only and require to perform in a quite unnatural task. Thus ERP studies allow to study prosodic processing more closely to normal natural speech.

Acknowledgements: We thank Katja Kühn, Annett Schirmer, and Franziska Kopp for their assistance in generating and analysing the signals. The work was supported by the Deutsche Forschungsgemeinschaft (FR 519/17-1).

7. REFERENCES

1. Alter, K., Steinhauer, K. & Friederici, A.D. De-accentuation: Linguistic environments and prosodic realizations. Paper presented at the ICSLP'98, Sydney, Australia (1998).
2. Cutler, A., Dahan, D. & van Donselaar, W. Prosody in the comprehension of spoken language: A literature review. *Language and Speech* **40**, 141-201 (1997)
3. Frazier, L. & Rayner, K. Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology* **14**, 178-210 (1982).
4. Kutas, M. & Hillyard, St. A. Reading senseless sentences: Brain potentials reflect semantic incongruency. *Science* **207**, 203-205 (1980).
5. Osterhout, L. & Holcomb, P. J. Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language* **31**, 785-804 (1992)
6. Steinhauer, K., Alter, K. & Friederici, A.D. Prosodic on-line effects on sentence processing: Parameters and ERP correlates. Paper presented at the CUNY 1998, New Jersey, USA (1998).
7. Warren, P., Grabe, E. & Nolan, F. Prosody, phonology, and parsing in closure ambiguities. *Language and Cognitive Processes* **10**, 457-486 (1995).