

Misleading dialogues: Human's brain reaction to prosodic information

Claudia Hruska*, Kai Alter*, Karsten Steinhauer[§] and Anita Steube[#]

*Max-Planck-Institute of Cognitive Neuroscience, Leipzig/Germany; [§] Georgetown University, Washington DC/USA; [#] University of Leipzig/Germany

Abstract

Fort peu de travaux ont eu pour but d'étudier les bases neurales et le décours temporel du traitement de la prosodie dans des situations communicatives comparables aux dialogues. Dans notre étude, nous analysons les réactions cérébrales provoquées lors de dialogues prosodiquement corrects et incorrects en utilisant la méthode des Potentiels Evoqués (PEs).

Nous parvenons aux deux résultats suivants. D'une part, il nous est possible de déterminer un reflet électrophysiologique du traitement du placement des accents, le système cognitive privilégiant le traitement des accents liés dans les réponses à des informations prépondérantes.

D'autre part, nous constatons que le traitement d'une réponse prosodiquement incorrecte est influencé par la position des accents, notamment dans le cas où l'accent déterminé par la question précédente n'apparaît pas dans la position prévue et présentant une négativité plus ample dans les PEs.

1 Introduction

In communication settings, humans use prosody to structure information. Prosody is an intrinsic determinant of spoken language (Cutler, Dahan & van Donselaar 1997), and prosodic cues are used on all levels of speech processing. As compared to written language, these prosodic features provide additional information about both the underlying syntactic structure and the information structure of the utterance (Beckman 1996). Empirical evidence for the usage of prosodic cues in early speech perception was reported by Steinhauer, Alter & Friederici (1999) for intonational phrase boundaries (IPh) in sentences presented in isolation. The authors used event-related brain potentials (ERPs) and identified a specific ERP correlate for the processing of IPh-boundaries, termed the Closure Positive Shift (CPS). The ERP technique is of particular advantage in the on-line study of speech perception. It has a very high time resolution, and ERP components reflect systematic brain activity related to both the perception of physical input and higher cognitive processes without requiring the interruption of the speech signal for data collection.

One important issue concerns the employment of prosodic information in daily conversation. In dialogues, it can often be observed that non-native speakers, even if almost fluent in the foreign language, do not appropriately realise the crucial prosodic speech parameters, resulting in severe misunderstandings. In the study at hand we investigated the role of correct and inappropriate accentuation in a very simple dialogue situation, namely question-answer-pairs. In terms of information structure, questions focus on expected *new* information. In the answer, this new information is then prosodically emphasised by an accent, whereas the background information is de-accented.

2 Methods

2.1 Participants and design

Twenty-one healthy right-handed subjects (mean age 23) were instructed to listen to question-answer pairs and to judge whether the answer had the appropriate prosody, given the context question.. Four conditions of question-answer pairs (4 x 96, see **Table 1**) were presented with either appropriate or

Table 1: Examples of the presented question-answer pairs of the four conditions. Accentuated constituents are marked by bold letters. The literal translation of the answer "Peter verspricht Anna zu arbeiten und das Büro zu putzen" is "Peter promises Anna to work and to clean the office".

Narrow focus on second noun	
Wem verspricht Peter zu arbeiten und das Büro zu putzen? (Whom does Peter promise to work and to clean the office?)	
appropriate prosodic pattern in answer sentences	A1 Peter verspricht Anna zu arbeiten und das Büro zu putzen.
inappropriate prosodic pattern in answer sentences	I1 Peter verspricht Anna zu arbeiten und das Büro zu putzen.
Narrow focus on second verb	
Was verspricht Peter Anna zu tun? (What does Peter promise Anna to do?)	
appropriate prosodic pattern in answer sentences	A2 Peter verspricht Anna zu arbeiten und das Büro zu putzen.
inappropriate prosodic pattern in answer sentences	I2 Peter verspricht Anna zu arbeiten und das Büro zu putzen.

inappropriate prosodic patterns in the answer sentences. Position of the focused (accented) constituent varied between NP2 ("Anna") and verb2 ("arbeiten"). All dialogues were presented in a pseudo-randomised order in 8 blocks of 48 trials, distributed over two experimental sessions.

2.2 Acoustic data

48 sentences per condition were produced by a male (questions) and a female (answers) speaker of standard German and recorded in a soundproof chamber. The female speaker produced the stimuli as response to the specific given question focus. The stimuli were digitised (44.1 kHz/16 bit sampling rate) and measurements of duration of constituents and fundamental frequency were carried out. These were subjected to statistical analysis.

2.3 EEG and Data Acquisition

Electroencephalogram (EEG) was recorded, trials containing eye blinks and movement artefacts were rejected. All channels were referenced against the left mastoid and offline re-referenced against linked mastoids. The ERPs were averaged per condition and for each subject. Figures display the *Grand average* ERPs for the whole group (N=21).

3 Results

3.1 Acoustic analyses

In intonational languages such as German, the fundamental frequency (F0) provides the main information for discrimination between given and new information. All 48 sentence stimuli (answers) in each of the two main conditions were acoustically analysed, critical parameters being the F0 maxima of the first and the second noun phrase, the second verb phrase and the highest value in the conjunction phrase. The results (**Fig.1**) emphasise significant differences between conditions on the second noun phrase [$t(190) = 24.75; p < .0001$] as well as on the second verb phrase [$t(190) = 39.77; p < .0001$].

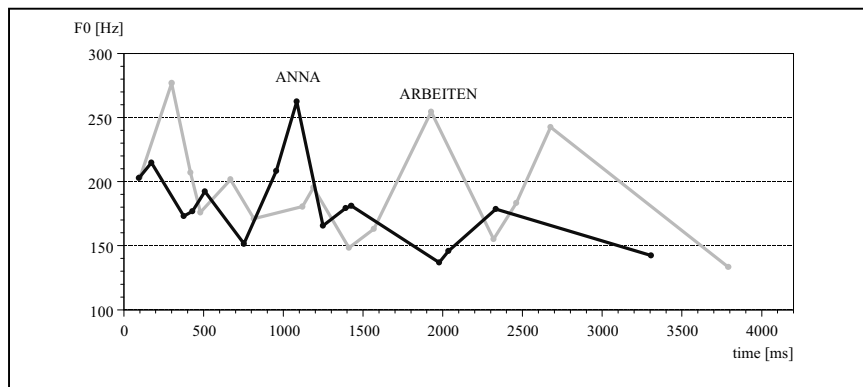


Figure 1: Pitch contour (based on F0 values) across the whole answer sentences due to the condition with focus on the second noun and of type A1 sentences (black line) and with focus on the second verb and of type A2 sentences (grey line).

Acoustically, also the duration of constituents differed between conditions. However, only the duration of the second verb differed significantly between conditions [$t(190) = 12.48$; $p < .0001$], whereas that of the second noun did not [$p > .05$]. Thus, in terms of perception of the accentuation, we assume the change in fundamental frequency as the more reliable parameter.

3.2 Behavioural data

The performance of the participants judging the prosodic appropriateness of the answer sentences was excellent (98% correct). There was no significant accuracy effect of condition, block, or session.

3.3 Psychophysiological ERP data

3.3.1 Brain responses to appropriate prosodic patterns (conditions A1, A2)

The event-related potentials in Figure 2 illustrate how different accent positions are processed online by the listeners. After receiving the new information in *noun-focused* sentences (A1), a large positive going waveform, the Closure Positive Shift (CPS), occurred early. When the *second verb* was in focus (A2), the CPS shifted towards the offset of this verb.

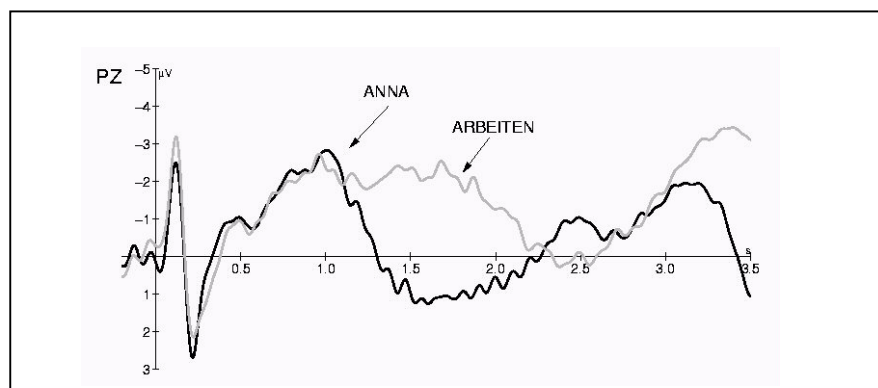


Figure 2: Grand averages due to appropriate prosodic pattern in answer sentences plotted here over the whole sentences for the condition with focus on the second noun (A1; black line) and second verb (A2; grey line).

3.3.2 Brain responses to inappropriate prosodic patterns (conditions I1, I2)

Violation of the information structure leads to widespread effects in the ERP. Figure 3 displays the grand averages for both accent positions in direct comparison to the same acoustic input once with appropriate and once with inappropriate prosodic patterns. Depending on the preceding context question, the same acoustic input leads to completely different responses.

When a *noun*-accentuated sentence was presented following a *verb*-focused question (Figure 3, left panel), the ERP differed considerably at the position where the new information was expected compared to the appropriate answer patterns due to noun focus questions. The missing accent at the

verb position elicited a large posterior negativity. Interestingly, the superfluous sentence accent on the second noun did not cause a mismatch effect.

Figure 3 (right panel) illustrates the processing of verb-accentuated sentences. This opposite kind of violation requires the mismatch detection for focus on the second noun because the second verb „arbeiten“ carries the accent. The missing accent on „Anna“ elicited a large negativity. This time, it was followed by a large positivity which resembled the CPS.

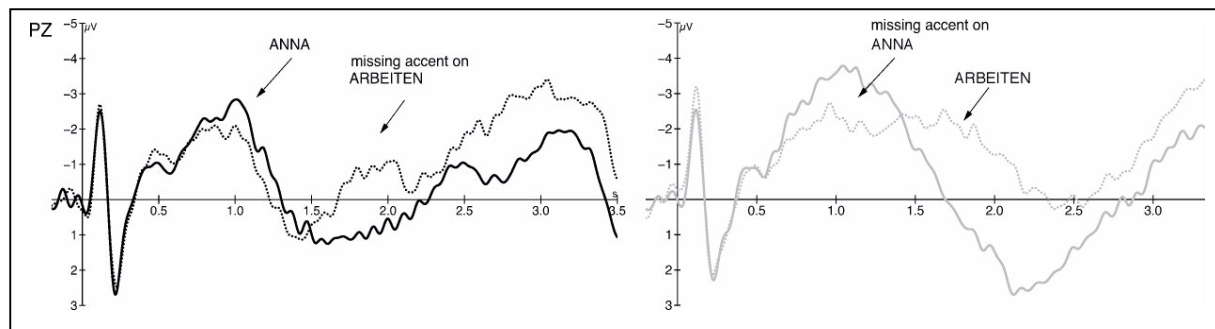


Figure 3: Grand averages at PZ electrode due to appropriate prosodic patterns and violations to the focus in preceding questions. **(a)** The left panel shows the ERPs of noun-accentuated sentences, once as the correct answer (A1; black line) and once as the inappropriate pattern due to verb focus (I2; dotted black line). **(b)** The right panel shows responses due to verb-focused sentences, again once for correct accentuation pattern (A2; grey line) and once due to noun focus in question (I1; dotted grey line).

4 Discussion

Exhaustive analyses of speech data illustrate how, in native German speakers, focus leads to realisation of characteristic prosodic patterns in the answer sentences of spoken dialogues.

The high accuracy of the behavioural data indicates that the prosodic appropriateness of the question-answer-pairs was easy to judge. A post hoc questionnaire revealed that participants used the rise and fall of the voice (F0) for identifying the appropriateness of accentuation. Taken together with a acoustic analyses of the speech signals these data demonstrate that fundamental frequency seems to be the most reliable indicator for accentuation in spoken dialogues.

ERP measurements show that focused information guides the perception of answer sentences in dialogues. The new information in these answers is the most important factor for the listeners. The Closure Positive Shift (CPS) following focused constituents suggests that focused elements are critical for phonological phrasing.

Inappropriate prosodic patterns are indicated by positions where the new information is expected to be emphasised through accent. Missing accents lead to a large negativity in the ERP, whereas superfluous accentuation seems to be a less salient cue for the listeners. These findings indicate that missing rather than superfluous accents trigger mismatch detection in dialogues.

5 Conclusion

This study combined acoustic analyses of speech data with the on-line ERP measurement during perception of spoken dialogues in German for the first time. The data show that context information provided by questions has a main influence on the processing of answers. Listeners pay attention mainly to the focused constituents in the answer which contain the expected and salient new information. Unexpected prosodic cueing, namely missing accentuation, seems to make it difficult to extract the relevant information from an already given context.

References

1. Beckmann, M. E. 1996. "The parsing of prosody" *Language and Cognitive Processes*, 1, 17-67.
2. 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.
3. Hruska, C., Alter, K., Steinhauer, K. & Steube, A. 2000. "Can wrong prosodic information be mistaken by the brain?" *Journal of Cognitive Neuroscience*, Suppl., 122.
4. Steinhauer, K., Alter, K. and Friederici, A.D. 1999. "Brain responses indicate immediate use of prosodic cues in natural speech processing" *Nature Neuroscience*, 2, 191-196.

Acknowledgement

The work was supported by the Leibniz Science Prize awarded to A.D. Friederici and a grant from the German Research Foundation DFG FR 519/ 17-2.