Segregating early physical and syntactic processes in auditory sentence comprehension

Anja Hahne,CA Erich Schröger[1] and Angela D. Friederici

Max Planck Institute of Cognitive Neuroscience, PO Box 500355, D-04303 Leipzig; [1]University of Leipzig, Leipzig, Germany

CA Corresponding Author

Received 26 November 2001; accepted 19 December 2001

Auditory language comprehension involves physical as well as syntactic processing. The present study examined whether early physical and syntactic processes in spoken sentence comprehension can be segregated using event-related brain potentials (ERPs). In the physical manipulation condition, the terminal word of the sentence was presented either from the same or from a different location to the preceding sentence fragment. In the syntactic manipulation condition, the terminal word was either a syntactically correct continuation of the preceding sentence fragment or violated syntactic constraints. These two factors were completely crossed. Physical deviances elicited the mismatch negativity (MMN) and syntactic deviances the early syntax-related negativity, both deviance-related components of the ERP. Sentences which violated physical as well as syntactic constraints elicited a negativity which was larger than that elicited by only a physical or only a syntactic deviance. The elicitation of the MMN in connected speech demonstrates that this component can be used as a probe for auditory change-detection even in ecologically highly valid situations. The increase of deviance-related effects with double deviants suggests that the early physical and syntactic processing systems act, to a high degree, in parallel and independently of each other. *NeuroReport* 13:305–309 © 2002 Lippincott Williams & Wilkins.

Key words: Auditory event-related potentials; Language; Mismatch negativity; Syntax

INTRODUCTION

When auditory information is processed the input develops sequentially over time. In order to effectively process this informational stream, our brain segments it into appropriate units such as phonemes or words (with speech input) or other perceptual objects (with physical input). Moreover, it generates expectations about the forthcoming input which, for example, enable fast categorization.

In the domain of physical feature processing, a vast amount of research has examined how changes in a repetitive auditory environment are processed. The measuring of event-related potentials (ERPs) revealed a characteristic component which is sensitive to irregularities in auditory stimulation [1,2], called mismatch negativity (MMN). This component is fronto-centrally distributed and peaks around 100–250 ms after the onset of the deviant stimulus. The MMN is a preattentive brain response which is thought to reflect the result of a comparison process registrating the deviance of the current auditory stimulus from the sensory memory trace of a previously processed reference stimulus. In the domain of language, a negativity with a similar latency as the MMN has been observed. This component, which is thought to reflect the processing of syntactic deviances, has been observed in correlation with violations of the phrase structure of a sentence. It has been observed with visual [3,4] as well as with auditory presentation [5–8]. In these studies, the component was observed on a critical word whose syntactic category did not match the categorical constraints of the prior sentence context. This syntax-related negativity has been interpreted as a reflection of first-pass parsing processes. Typically, it is followed by a late centro-parietally distributed positivity (P600) which is taken to reflect processes of syntactic integration, reanalysis or repair.

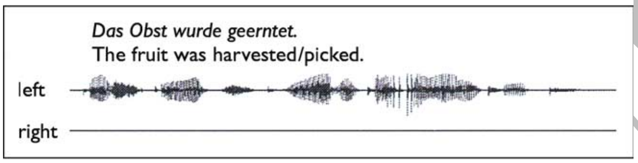
These two negativities, the MMN elicited by physical deviances and the early anterior negativity elicited by a syntactic violation, bear a variety of similarities with regard to polarity, latency and topography. The aim of the present study was to determine whether these two kinds of detection systems, i.e. the physical and the syntactic system, are distinct or not. We explored whether the physical system and the language system interact in the processing of physical and syntactic deviances or whether these deviance-detection systems operate independently of one other. To address this issue, we designed an experiment in which physical features and syntactic features were systematically manipulated. More specifically, participants listened to spoken sentences which were presented either continuously from a single location or included a location switch (physical manipulation) and which were either syntactically correct or not (syntactic manipulation). The syntactic deviance was realized as a phrase structure violation whereas the physical deviance was realized as a location switch, i.e. the start of a sentence was presented from a loudspeaker on the left (or right) while the final word of the sentence was presented from a loudspeaker on the right (left). These two factors, syntax and location, were completely crossed resulting in four different experimental conditions (Fig. 1). The participants' task was to judge the grammaticality of each sentence.

The comparison of conditions (a) and (b) in Fig. 1 allows us to examine the effect of pure physical deviance. We expected a MMN for this comparison. The comparison of conditions (a) and (c) allows the examination of the pure syntax effect. We expected to observe an early anterior negativity for this comparison. Condition (d) combines physical with syntactic deviance and is therefore most critical with regard to the question at hand. A direct comparison of the difference between conditions (a) and (d) on the one hand and the summation of the difference between conditions (a) and (b) and the difference between condition (a) and (c) on the other hand, will reveal whether the two deviating features, syntax and physical location, elicit fully additive effects or whether they interact.

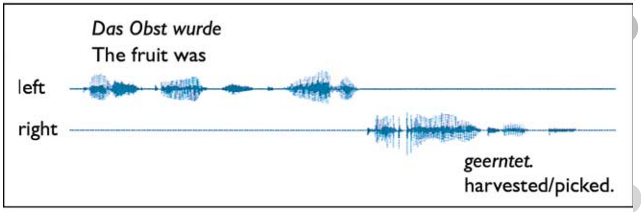
MATERIALS AND METHODS

Twenty-four paid right-handed native speakers of German participated in the experiment (age 18–27 years, mean age 23; 10 male).

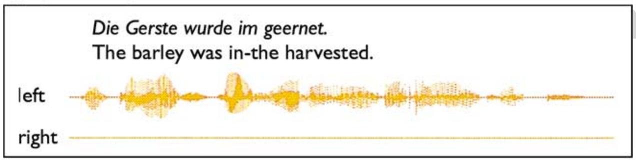
(a) Syntactically correct sentence without a location switch



(b) Syntactically correct sentence with location switch



(c) Syntactically incorrect sentence without a location switch



(d) Syntactically incorrect sentence with location switch

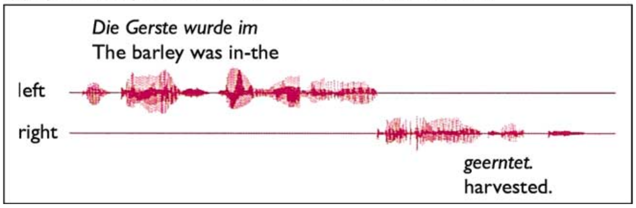


Fig. 1. Illustration of the four experimental conditions. Sentences were either presented from one location or contained a location switch for the critical terminal word. In addition, sentences were either syntactically correct or contained a phrase structure violation. Approximate literal translations are provided.

To create the critical sentences, 96 verbs were selected. With these verbs, 96 correct sentences of the form 'determiner–noun–auxiliary–participle' (Fig. 1), and 96 syntactically incorrect sentences of the form 'determiner–noun–auxiliary–preposition–participle' were constructed. In the incorrect sentences a preposition was directly followed by a verb form which created a phrase structure violation realized as a word category error. In addition, two types of filler sentences were included: first, 96 correct sentences in which a preposition was followed by a noun. This condition was included to ensure that participants could not already predict the syntactic incorrectness of the sentence on the preposition. Second, 96 additional syntactically incorrect sentences with different participles were created in order to equate the probability of correct and incorrect sentences within the experiment. Thus, there were 384 different sentences.

All sentences were spoken by a trained female native speaker of German. The sentences were recorded on digital audio tape and then sampled at 20 kHz with a 16 bit resolution. Each sentence was stored as a separate file on the hard disk of the computer for presentation during the experiment. To enable a precise time locking of the ERP, the onset of the participle was marked. In creating the syntactically incorrect sentences, we followed the procedure described in Hahne and Friederici [7]. Four different

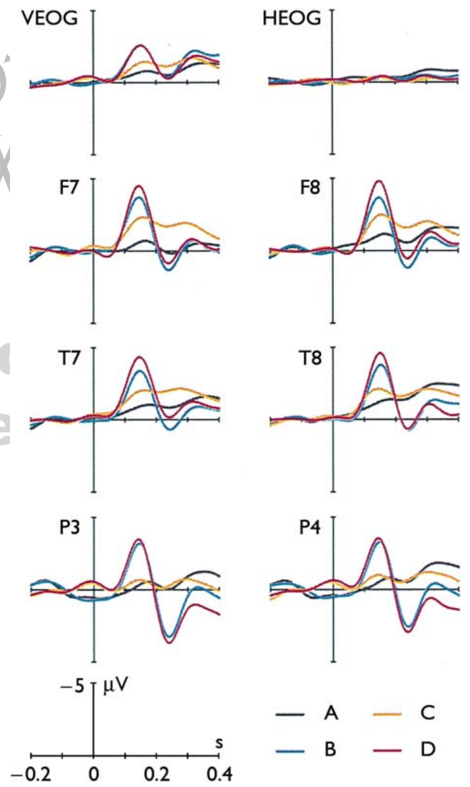


Fig. 2. Grand average ERPs on the critical word (participle) elicited by correct sentences with (light blue) or without location switch (dark blue), syntactically incorrect sentences with (magenta) or without location switch (orange) for six electrode positions. The waveforms in this figure were smoothed by a low-pass filter with a cut-off frequency (~3 dB) at 10 Hz.

versions of each sentence were created: (a) the complete sentence on the left channel, (b) the complete sentence on the right channel, (c) the participle on the left channel with the previous sentence on the right channel and (d) the participle on the right channel with the previous sentence on the left channel.

The 96 sentences of each of the four sentence types (two critical and two filler types) were systematically distributed over six experimental lists using a Latin square design such that each item appeared equally often as a standard stimulus (i.e. without location switch) and as a deviant stimulus (i.e. with a location switch) and that each list contained 48 items of each sentence type and an equal number of standard and deviant trials. These six lists were individually pseudo-randomized with the constraints that no more than four standard *vs* deviant trials and no more than three identical sentence types and no more than four syntactically incorrect sentences followed in direct succession. Based on these six lists, another six lists were created by exchanging standard and deviant trials. These 12 lists were realized with the standard on the left as well as with the standard on the right, thus resulting in 24 different presentation lists, each of which was realized exactly once in the experiment.

Participants were seated in a comfortable chair in a sound-attenuated booth, 100 cm in front of a computer screen. At a distance of 155 cm, two loudspeakers were placed at head-level. The angle between the participant and the two speakers was 39°. Each trial started with a fixation signal, which appeared on the screen 500 ms before the sentence presentation started and remained visible until it was replaced by a response sign 1500 ms after sentence offset. The response cue was presented for maximally 2 s but disappeared on response. The next trial started after an inter-trial interval of 1500 ms. Participants were asked to avoid blinks or other movements during the presentation of the fixation signal and to judge the sentences for grammatical correctness by pressing one of two buttons with their thumbs during the presentation of the response signal. Participants were told that the sentences would usually be presented via the left (right) speaker but sometimes also via the right (left) speaker but that their task was to judge the grammaticality of the sentences for which changes in location would be irrelevant. Sentences were presented in four blocks containing 96 trials each. Prior to the experimental trials, 18 practice sentences were presented.

The EEG was recorded from 27 Ag/AgCl electrodes mounted in an elastic cap. The vertical electrooculogram (VEOG) was recorded from electrodes placed above and below the right eye. The horizontal EOG (HEOG) was recorded from electrodes at the outer canthus of each eye. All recordings were referenced to the left mastoid. Electrode impedance was kept below 5 kOhm. EEG and EOG were amplified with a DC amplifier and digitized with 250 Hz.

Error rates were computed separately for each condition. ERPs were averaged only for correctly answered trials. Trials in which participants had given a wrong or no grammaticality judgment response and trials contaminated by eye movements or amplifier blocking were discarded. These were distributed equally across conditions. ERPs were calculated for 400 ms beginning with the presentation of the participle relative to a 200 ms prestimulus baseline.

Analyses on mean voltage were performed within the latency window of 125–175 ms, which corresponds to a 50 ms interval around the peak latency of the MMN at Fz (144 ms). Repeated measure ANOVAS were performed with four within-subject variables for the lateral electrodes: location (standard *vs* deviant), syntax (correct *vs* incorrect), hemisphere (left *vs* right) and region (anterior *vs* central *vs* posterior). The variables hemisphere and region were completely crossed yielding six regions of interest: left anterior (F7,F3,FT7,FC3), right anterior (F8,F4,FT8,FC4), left central (T7,C3,TP7,CP5), right central (T8,C4,TP8,CP6), left posterior (P7,P3,O1), right posterior (P8,P4,O2). The Geisser–Greenhouse correction was applied when necessary.

RESULTS

Behavioral data: Error rates were significantly higher for syntactically incorrect sentences (1.8%) than for syntactically correct sentences (0.6%; main effect of syntax: F(1,23) = 15.87, $p < 0.001$, MSE = 0.51). At the same time, reaction times were reliably faster for syntactically incorrect sentences (354 ms) than for syntactically correct sentences (376 ms; main effect of syntax: F(1,23) = 4.88, $p < 0.05$, MSE = 2407.85). However, these significant effects are difficult to interpret as they may reflect a speed–accuracy trade off rather than a genuine effect.

The main effect of location was not significant in either analysis (error: F < 1; reaction time: F(1,23) = 1.72, $p < 0.20$, MSE = 524.20) nor was the interaction of location and syntax (error and reaction time: F < 1).

ERP data: The ERP data for the critical word in all four conditions are displayed in Fig. 2.

As predicted, correct sentences which switched their location elicited a MMN in comparison to correct sentences which did not switch their location. Furthermore, when comparing sentences which did not switch their location, syntactically incorrect sentences elicited an early anterior-central negativity relative to correct sentences. Interestingly, the syntactically incorrect condition which switched its location elicited an even larger negativity than the correct condition which switched its location.

The ANOVA revealed significant main effects of location (F(1,23) = 130.46, MSE = 8.90, $p < 0.0001$) and syntax (F(1,23) = 9.14, MSE=6.72, $p < 0.01$). However, there was no interaction of location and syntax (F < 1). Location as well as syntax interacted significantly with the variable region (location: F(2,46) = 39.50, MSE = 1.22, $p < 0.0001$; syntax: F(2,46) = 18.38, MSE = 1.07, $p < 0.0001$). No other interaction reached significance.

When analysing the three levels of the variable region separately, highly significant main effects of location and syntax were obtained for the anterior (location: F(1,23) = 129.36, MSE = 4.84, $p < 0.0001$; syntax: (F(1,23) = 36.92, MSE = 2.09, $p < 0.0001$) and the central region (location: F(1,23) = 147.76, MSE = 3.32, $p < 0.0001$; syntax: F(1,23) = 8.69, MSE = 2.72, $p < 0.01$). For the posterior region, the main effect of location was also highly significant (F(1,23) = 44.17, MSE = 3.17, $p < 0.0001$) whereas there was no significant main effect of syntax (F < 1). There was no reliable interaction between location and syntax in either region (anterior: F(1,23) = 1.74, MSE = 2.67, $p < 0.20$; central and posterior: F < 1).

Although the global analysis did not reveal any interaction of location and syntax, a visual inspection of the data suggests that the two variables were not completely independent. As the nonsignificant interaction might be a result of insufficient power, we analyzed whether the syntactically incorrect condition with a location switch elicited a statistically larger deviance-related negativity than the syntactically correct condition which switched the location. This comparison revealed a significant interaction of condition and region (F(2,46) = 8.84, MSE = 0.76, $p < 0.01$). The condition effect was significant for anterior electrode positions (F(1,23) = 5.93, MSE = 3.70, $p < 0.05$), but not for central (F(1,23) = 2.66, MSE = 3.23, $p < 0.12$) and posterior positions (F < 1). This indicates that at anterior electrode sites, the negativity for the deviant stimuli that differed on both dimensions, location and syntax, was larger than the negativity for the stimuli differing in location only.

The fact that the variables location and syntax did not reveal any reliable interactions suggests that the two processes reflected in these effects are rather independent. In order to further test whether the MMN and the syntax-related negativity indeed elicited independent effects, we conducted additional analyses comparing the arithmetic summation of the MMN (location switch, correct sentence) and the syntax-related negativity (no location switch, incorrect sentence) to the empirically observed combined condition (location switch, incorrect sentence). These analyses revealed a significant main effect of condition (F(1,23) = 8.34, MSE = 10.99, $p < 0.01$) as well as a reliable interaction of condition and region (F(1,23) = 7.20, MSE = 1.06, $p < 0.01$). Analyses for each level of region revealed reliable effects for condition for the anterior (F(1,23) = 10.41, MSE = 6.50, $p < 0.01$) and the central region (F(1,23) = 7.77, MSE = 4.10, $p < 0.05$), but not for the posterior region (F(1,23) = 2.92, MSE = 2.51, $p < 0.11$), the sum being larger than the double-deviance effect.

DISCUSSION

The present study examined the relationship between language-related features and physical features in auditory sentence processing. The terminal word of a sentence which was either a correct or a syntactically incorrect continuation was presented either from the identical location as the sentence fragment or contained a location switch. Correct sentences which switched their location showed a MMN compared to correct sentences which did not switch their location. This finding basically replicates previous results on MMN and location changes [9–16]. However, previous studies showing a MMN for location changes always used a repetitive stimulation of discrete auditory events. Thus, the present work adds to these results by demonstrating that a location change within a continuous speech stream is also able to elicit MMN. This demonstrates that the MMN can be used as a probe for the pre-attentive auditory change detection even in ecologically more valid situations.

The comparison of sentences which did not switch their location but which were either syntactically correct or incorrect revealed an early anterior-central negativity. This again replicates previous results [5–8]. The critical question of the present study was whether these two processes, the physical feature analysis on the one hand and the syntactic feature analysis on the other hand, act independently of each other or not. A possible assumption is a temporal dependency with the comparison for the physical feature necessarily being completed before the syntactic comparison can be performed. The finding that stimuli which contained both a physical and a syntactic mismatch clearly elicited a more negative response than stimuli which contained a mismatch on only one dimension rules out the assumption that only one of these two deviance-detection processes can be active at a time. By contrast, the present data suggest that there is a vast amount of autonomy of physical and syntactic feature processing in this early stage of analysis. Statistical analyses did not reveal any significant interaction between the physical and the syntactic variable suggesting that physical and syntactic features of an auditory language stimulus are processed in parallel and independently to a large extent.

However, the observed additivity was not perfect. A direct statistical comparison between the arithmetic summation of the MMN effect in the physical condition and the syntax-related negativity in the language condition differed significantly from the empirically observed effect in the combined physical and syntactic mismatch condition. One possible explanation for this result is that the two processes may in fact not be completely independent of each other but share some overlapping resources such as the attentional demands needed to perform the required judgment task. Indeed, it was demonstrated that the MMN consists of two subcomponents, one represented by bilaterally located temporal generators and one represented by a right inferior frontal source. The latter of these is taken to reflect attentional aspects of deviancy detection [17–20]. Generators underlying the syntax-related negativity are also located in temporal and frontal areas [21]. Similar to the MMN, there are bilateral temporal generators, but unlike the MMN, the syntax-related negativity involves a left and a right frontal source. Structurally, the right frontal source is shown to involve the same brain region. It is, therefore, possible that the two components also overlap functionally with attentional aspects being the critical variable present in both.

CONCLUSION

Taken together, the present study demonstrates that during an early processing stage (i.e. within the first 200 ms) for an auditory speech stimulus, the physical deviance detection system and the language grammaticality system detecting syntactic violations can operate in parallel. In other words, the processing of syntactic deviances is independent of the processing of physical deviances occurring in the same acoustic input, at least to a high degree.

REFERENCES

1. Näätänen R. *Attention, and Brain Function*. Hillsdale, NJ: Erlbaum; 1992.
2. Schröger E. *Behav Res Meth Instr Commun* **30**, 131–145 (1998).
3. Neville HJ, Nicol J, Barss A *et al. J Cogn Neurosci* **3**, 151–165 (1991).
4. Gunter TC, Friederici AD and Hahne A. *Neuroreport* **10**, 3175–3178 (1999).
5. Friederici AD, Pfeifer E and Hahne A. *Cogn Brain Res* **1**, 183–192 (1993).

6. Hahne A. *J Psycholing Res* **30**, 251–266 (2001).
7. Hahne A and Friederici AD. *J Cogn Neurosci* **11**, 193–204 (1999).
8. Hahne A and Jescheniak JD. *Cogn Brain Res* **11**, 199–212 (2001).
9. Deouell LY and Bentin S. *Psychophysiology* **35**, 745–754 (1998).
10. Kaiser J, Lutzenberger W and Birbaumer N. *Neuroreport* **11**, 2889–2892 (2000).
11. Kaiser J, Lutzenberger W, Preissl H *et al*. *J Neurosci* **20**, 6631–6639 (2000).
12. Paavilainen P, Karlsson ML, Reinikainen K and Näätänen R. *Electroencephalogr Clin Neurophysiol* **73**, 129–141 (1989).
13. Schröger E. *Psychophysiology* **32**, 55–65 (1995).
14. Schröger E. *Hearing Res* **96**, 191–198 (1996).
15. Schröger E and Wolff C. *Neuroreport* **7**, 3005–3008 (1996).
16. Winkler I, Tervaniemi M, Schröger E *et al*. *Neurosci Lett* **242**, 49–52 (1998).
17. Giard MH, Perrin F, Pernier J and Bouchet P. *Psychophysiology* **27**, 627–640 (1990).
18. Näätänen R, and Alho K. *Brain Topogr* **7**, 315–320 (1995).
19. Rinne T, Alho K, Ilmoniemi RJ *et al*. *NeuroImage* **12**, 14–19 (2000).
20. Opitz B, Rinne T, Mecklinger A *et al*. *NeuroImage* (in press).
21. Friederici AD, Wang Y, Herrmann CS *et al*. *Hum Brain Mapp* **11**, 1–11 (2000).

Acknowledgements: We thank Ina Bornkessel and Bertram Opitz for helpful comments on a previous version of this manuscript, Valentin Wagner for his help in preparing the experiment and Comelia Schmidt for data collection.