Pitch accent type affects the N400 during referential processing

Petra B. Schumacher[a] and Stefan Baumann[b]

Discourse processing depends on semantic memory as well as maintaining and updating of a mental model. Using event-related potentials, we investigated how a referent's information status (new, accessible, given) is processed in combination with three different prosodic realizations (an appropriate accent and two inappropriate accents). The data reveal a biphasic N400-late positivity pattern, indicating that prosodic information affects an early discourse linking stage, during which prominence information reflecting a referent's accessibility is computed (N400), and a later discourse updating stage, during which conflicts between prosodic information and a referent's actual information status are detected (late positivity). Crucially, the data show that the N400 is not only sensitive to lexico-semantic relations but also to discourse accessibility induced by prosodic cues. *NeuroReport* **21:618–622 © 2010 Wolters Kluwer Health | Lippincott Williams & Wilkins.**

NeuroReport 2010, **21**:618–622

Keywords: cognitive activation, discourse, language processing, late positivity, N400, pitch accent, prosody, reference tracking, semantic memory

[a]Independent Emmy Noether Research Group, Department of English and Linguistics, Johannes Gutenberg University Mainz, Mainz and [b]Institute of Linguistics, Phonetics, University of Cologne, Cologne, Germany

Correspondence to Dr Petra B. Schumacher, Department of English and Linguistics, Johannes Gutenberg University Mainz, Jakob-Welder-Weg 18, 55099 Mainz, Germany
Tel: +49 6131 3923859; fax: +49 6131 3923836;
e-mail: petra.schumacher@uni-mainz.de

Received 22 February 2010 **accepted** 8 March 2010

Introduction

Information exchange between individuals generally takes place through utterances that express relations between referents in specific states or events. The individual constituents of an utterance can be regarded to have a certain information status or degree of cognitive activation (i.e. how active an entity is in the consciousness of a speaker and a hearer). Neuropsychological research in the visual domain (reading) has shown that a referent's information status affects different processing stages during the construction of a mental discourse model: semantic memory supports linking processes and can be dissociated from processes subserving the updating of the mental model [1,2]. On the basis of electrophysiological measures, a referent's information status is reflected in amplitudinal modulations of the N400, a negativity peaking at approximately 400 ms after the onset of a referential expression; discourse updating results in a later positive deflection. Crucially, the information status can be marked by morpho-syntactic cues (e.g. *a reindeer* introduces new information, whereas *the reindeer* indicates a certain degree of familiarity with the referent), distinct lexical forms [e.g. a full noun phrase (NP) such as *the reindeer* vs. a pronoun like *it*, which points to a referent introduced earlier], but also by prosodic cues, such that different types of pitch accent indicate differences in the information status.

Information status is here understood to reflect distinct levels of cognitive activation for a referent, and we distinguish three levels (following Chafe [3]): (i) if a referent is already active in the listener's consciousness at the time of the utterance, it is *given*; (ii) if a referent becomes activated from a previously inactive state, it is *new*; and (iii) if a referent becomes activated from a previously semi-active state, it is *accessible*. This research focuses on the processing of a particular type of accessibility, in which a referent is accessible through inference from another active discourse referent, as in example (1) where *the sole* is accessible through a whole-to-part relation from the (active) expression *an old shoe* introduced earlier.

(1) Sabine repariert einen alten Schuh. Dabei zerschneidet sie **die Sohle**.
Sabine repairs an old shoe. In doing so, she cuts **the sole**.

Event-related potential (ERP) research on reading comprehension of inferentially accessible referents has identified a biphasic N400-late positivity pattern relative to the comprehension of already given, active referents [1,2,4]. In these studies, the enhanced N400 reflected difficulties during linking with the discourse model, and its amplitude increased with decreasing referent accessibility, where semantic and syntactic cues of referent accessibility were varied. This investigation extends these findings to the auditory (listening) domain to examine how *prosodic* cues (i.e. different accent types) are used to determine a referent's information status during language comprehension. The late positivity found in the reading studies was observed for costs from discourse updating and restructuring, for example, when an accessible but previously unmentioned discourse referent had to be introduced

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Website (*www.neuroreport.com*).

0959-4965 © 2010 Wolters Kluwer Health | Lippincott Williams & Wilkins DOI: 10.1097/WNR.0b013e328339874a

into the mental model. In accordance with these effects of referential interpretation, research in the auditory domain on focus prosody has reported N400-effects for linking problems, if expected prosodic cues (accents) were missing in the input, and late positivities for the reorganization of the discourse structure because of mismatches between discourse and prosodic information [5–7].

This study examined effects from differences in the prosodic marking of information status in German. For (West Germanic) languages, such as English, German, and Dutch, it is commonly assumed that new referents are marked by pitch accents, whereas given referents are deaccented [8,9]. With respect to the prosodic realization of *accessible* referents, however, several studies (such as [10] for American English and [11] for German) have shown that a simple dichotomy of accentuation versus deaccentuation is inappropriate for an account of accessible information, and that the *accent type* is an important cue for encoding a referent's information status. Perception and production experiments in German [12,13] showed that accessible information cannot be treated as a uniform category and that different types of activated information (part-to-whole, whole-to-part, hypernym-to-hyponym relations, etc.) demand different accent types as linguistic markers. In fact, for specific types of inferentially accessible information (relations inferred from a given scenario and whole-to-part relations), an early peak accent (H + L*) proved to be particularly appropriate in the perception study, as opposed to two other accent types that were tested, namely H* and deaccentuation (for new and given referents, respectively). [The notation of accent types follows the German ToBI (Tones and Break Indices) system (see Grice *et al.* [14]). 'H + L*' stands for a high falling accent with a low target on the accented syllable; 'H*' indicates a pitch peak on the accented syllable].

In this research, we wanted to investigate how prosodic cues influence the comprehension of inferentially accessible referents in the German language. We recorded the electrophysiological responses to referring expressions that can be linked to prior discourse through a whole-to-part relation (e.g. *shoe–sole*) using three different types of accentuation (following the perception study reported by Baumann and Grice [12]). We hypothesized that the language processor is sensitive to these prosodic distinctions. In particular – following previous purely auditory perception studies – we not only expected clear differences in the processing of accented versus deaccented items, but also modulations with respect to the appropriateness of different pitch accent types. Specifically, the accent type H + L* is hypothesized to be the most appropriate marker of the information status 'accessible' encoded by a whole-to-part relation. In turn, H* (marking new information) should be less acceptable, and deaccented referents should yield the most pronounced mismatch. Regarding prosodic cues as one means to encode a referent's information status, we predict N400 differences for linking costs, and possibly late positivities for the mismatch between prosodic information and the referent's actual information status.

Methods

Twenty-four right-handed, monolingual native speakers of German (12 men of 22–29 years of age, mean = 24.9 years) from the University of Mainz participated in the study after giving written informed consent. Three of them had to be discarded from the data analysis because of excessive ocular artifacts. Ninety sets of stimuli were constructed that consisted of two successive sentences each [see (1) above]: a context sentence, which included an anchor expression to make the critical NP inferentially accessible (e.g. in 'Sabine repairs an old shoe', *an old shoe* serves as an anchor for *the sole*), and a target sentence, which contained the critical NP (e.g. 'In doing so, she cuts the sole.'). Anchor expression and critical NP (*the sole*) always stood in a whole-to-part relation to each other. The syntactic and rhythmic structures were kept constant for all sets.

The stimuli were read by a trained phonetician and digitally recorded in a soundproof room. While the intonation pattern of the context sentence was not varied, three different versions of each target sentence were created. The critical NP either carried a nuclear H* ('newness accent') or H + L* pitch accent ('accessibility accent'), or was deaccented (with a nuclear L + H* assigned to the preceding verb) ('givenness accent'). [See Supplemental digital content 1, *http://links.lww.com/WNR/A45* for screen shots of the intonation patterns of the example in (1) and associated sound files for the three conditions, Supplemental digital content 2, *http://links.lww.com/WNR/A46*: context sentence, Supplemental digital content 3, *http://links.lww.com/WNR/A47*: target sentence with accessibility accent, Supplemental digital content 4, *http://links.lww.com/WNR/A48*: target sentence with newness accent, Supplemental digital content 5, *http://links.lww.com/WNR/A49*: target sentence with givenness accent]. To guarantee equivalent pitch accent types in all target sentences, we resynthesized the naturally produced contours whenever necessary, using the speech analysis tool PRAAT [15]. The resulting 270 critical trials were distributed across three lists (with 30 trials per prosodic realization each) and interspersed with 160 fillers. A comprehension question was created for each trial to assess the listeners' attention, for instance for stimulus (1) participants had to respond to the question 'Does Sabine cut the sole?' by pressing a 'yes' or 'no' button on a response box. 'Yes' and 'no' responses were equally distributed across all trials, and incorrect or timed-out responses were excluded from the ERP analyses. Responses to this task were at ceiling level and registered no differences between conditions (all Fs < 1).

Each session started with a brief practice block to familiarize participants with the procedure. The experimental session contained 250 pseudo-randomized trials

and was conducted in five blocks with pauses in-between. After the presentation of a fixation cross in the center of a screen for 500 ms, a trial was presented auditorily while the fixation cross remained on the screen. Following a 500 ms blank screen, the comprehension question was presented visually in the center of the monitor in yellow letters on a blue background. Participants were instructed to respond to this question as quickly and as accurately as possible by pressing a button on a response box, and response times were restricted to 4000 ms. After an intertrial interval of 1000 ms, the next trial started.

The electroencephalogram was recorded from 24 Ag/AgCl electrodes mounted in an elastic cap (*EasyCap*). Recordings were referenced to the left mastoid and rereferenced offline to linked mastoids (ground: AFz). The electrooculogram was recorded by electrode pairs placed above and below the participant's left eye and at the outer canthus of each eye. Impedances were kept below 5 kΩ. All channels were amplified using a *BrainVision Brain-Amp* amplifier and recorded with a digitization rate of 500 Hz. Data were bandpass-filtered offline (0.3–20 Hz) to exclude slow drifts that could lead to stimulus-independent differences. The ERP analysis included trials with correct responses to the comprehension question and without ocular or other artifacts. Average ERPs were time-locked to the onset of the critical NP (e.g. *the sole*) and computed per condition and participant, prior to grand-averaging over all participants. Repeated-measures analyses of variance were carried out with the factor PROSODY (three levels) and the topographical factor region of interest (ROI) for lateral [four levels: left anterior (F3/F7/FC1/FC5), right anterior (F4/F8/FC2/FC6), left posterior (P3/P7/CP1/CP5), right posterior (P4/P8/CP2/CP6)], and midline electrodes separately (six levels: Fz/FCz/Cz/CPz/Pz/POz). Analyses were carried out hierarchically on the mean amplitude value per condition and Huynh–Feldt correction was applied whenever the degree of freedom in the numerator was greater than one [16]. Pairwise comparisons were also subjected to repeated-measures analyses of variance. Since auditorily presented stimuli typically show earlier onset latencies than visually presented material (on which previous findings were based on), we adopted the following routine to determine the critical time windows: analyses were first carried out on successive 40 ms-epochs starting at the onset of the critical NP; at least three adjacent epochs with significant effects were required before being subjected to comprehensive statistical analyses. This yielded the time-windows from 320 to 440 ms and 520 to 640 ms.

Results

Figure 1 shows a biphasic pattern for the grand-averages of the three conditions. It reveals a three-way modulation of the N400 amplitude (givenness accent > newness accent > accessibility accent) followed by a more pronounced positivity for deaccentuation (i.e. givenness accent). Statistical analyses for the time window from 320 to 440 ms showed a main effect of PROSODY [over midline electrodes: $F(2,40) = 24.46$, $p < 0.001$, and over lateral sites: $F(2,40) = 21.21$, $p < 0.001$]. Pairwise comparisons by PROSODY in this time window showed reliable differences for all three contrasts: accessibility versus givenness accent [mid(line): $F(1,20) = 38.83$, $p < 0.001$; lat(eral): $F(1,20) = 29.23$, $p < 0.001$], newness versus givenness accent [mid: $F(1,20) = 30.8$, $p < 0.001$; lat: $F(1,20) = 51.43$, $p < 0.001$], and accessibility versus newness [midline only: $F(1,20) = 4.72$, $p < 0.05$; lat: $F < 1$].

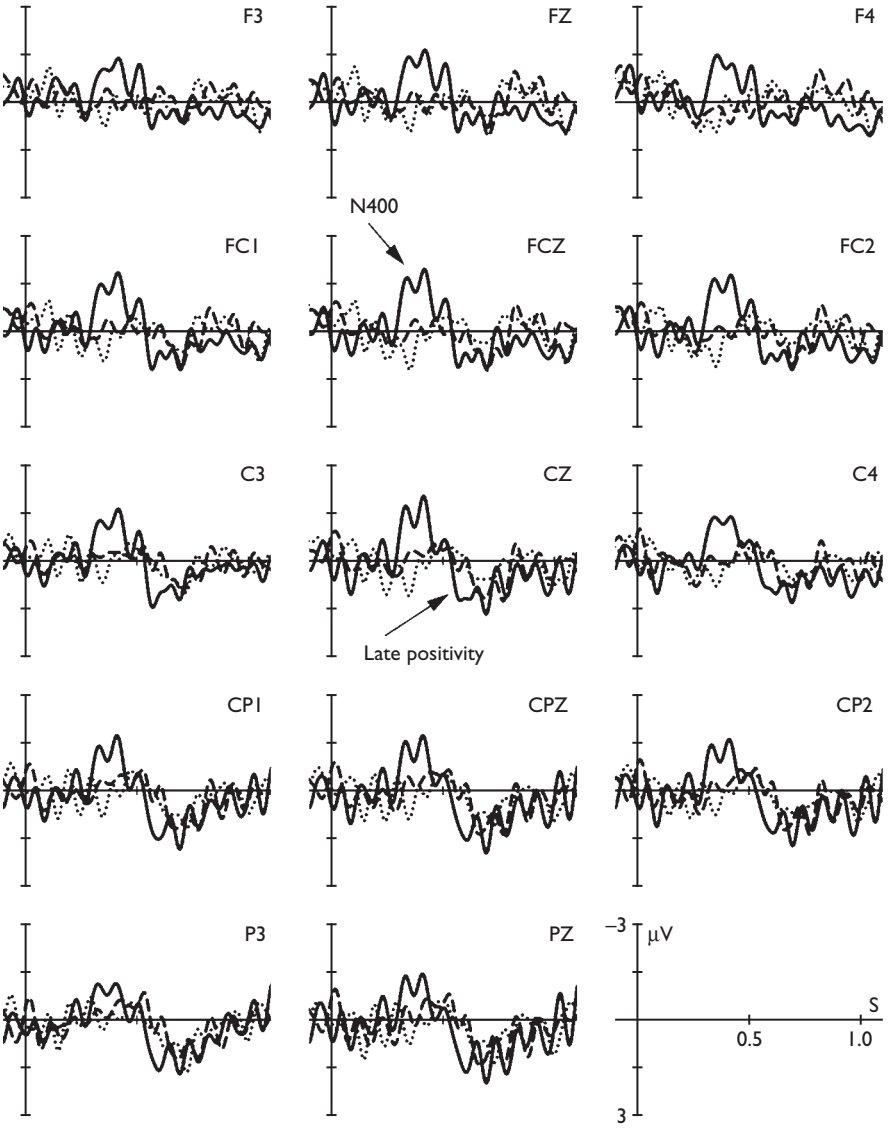
Analysis for the 520–640 ms interval registered a main effect of PROSODY [mid: $F(2,40) = 4.12$, $p < 0.03$, lat: $F(2,40) = 4.46$, $p < 0.02$] and an interaction of PROSODY × ROI for the analysis over lateral electrodes [$F(6,120) = 2.64$, $p < 0.04$], which was resolved over the two posterior ROIs. The main effect is reflected by a more pronounced positivity for givenness accent, which is supported by pairwise comparisons that showed significant main effects for givenness versus newness accent [mid: $F(1,20) = 9.66$, $p < 0.01$, lat: $F(1,20) = 9.17$, $p < 0.01$] and givenness versus accessibility accent [mid: $F(1,20) = 5.39$, $p < 0.03$, lat: $F(1,20) = 4.93$, $p < 0.04$], and no reliable differences between accessibility and newness accent (Fs < 1).

Discussion

The current investigation is the first ERP study targeting prosodic influences on referential accessibility proper, that is, in the absence of focus marking [5–7]. Previous research on the prosodic realization of different types of information status has already suggested a three-way distinction between accent types [12], but neuropsychological evidence for this distinction is provided here for the first time. Concerning the type of inferentially accessible information examined here, an early peak accent (H + L*) was identified to be the most appropriate prosodic realization, and the other two, less acceptable accent types (H* and deaccentuation) were hence predicted to show increased processing costs. Indeed, the ERP data show N400 differences as a function of accent type, with the most pronounced negativity for givenness accent and a somewhat enhanced negativity for newness accent. In addition, givenness accent registered a later positive deflection relative to the other two accent types. These findings converge with the research on referential processing in the visual domain and support the neurocognitive model laid out in [1,2,4], where referential integration proceeds in two steps: linking and discourse updating, reflected in N400 and late positivity modulations, respectively.

Regarding the N400, prominence relations are computed during discourse linking that reflect the referent's accessibility and support its integration with prior discourse: the more difficult the integration, the more enhanced is the N400 amplitude. Accent type has been taken as a specific prominence feature, and the current data indicate that this feature modulates the N400. The

Fig. 1



Grand average ERPs at 14 selected electrodes from 100 ms before until 1100 ms after the critical noun phrase (NP) (onset at vertical bar). Negativity is plotted upwards. Grand averages are filtered with 8 Hz low pass for the plot only. The expected 'accessibility accent' is shown in dotted lines, the 'newness accent' in dashed lines, and the 'givenness accent' in solid lines.

data thus show that prosodic information guides the computation of a referent's accessibility and can result in integration costs when less appropriate accent types are encountered. Crucially, in light of the fact that semantic relations were not manipulated, this indicates that the N400, which has traditionally been associated with lexico-semantic processing [17], is also sensitive to discourse-level information (here, accessibility encoded by prosodic cues). While it has already been shown that the N400 is not specific to language processing (e.g. arithmetics [18] or visual event processing [19]), the current findings thus suggest that the N400 must be more generally associated with prominence relations, rather than with semantic relations alone; that is, a referent's accessibility is determined by prominence features such as accent type or syntactic position. Research on (de)accentuation has further reported another negativity peaking at 400 ms with a frontal maximum, which is observed when a predictable accent is encountered (the so-called Expectancy Negativity) [5,20]. We do not believe that the present negativity is an Expectancy Negativity, because it has a broad distribution and the accent types were not predictable from prior context; compare [5].

With regard to the late positivity, information from different (linguistic) domains is assessed during discourse updating, and conflicting information result in the reorganization of the discourse model [6,7]. In the current case, a mismatch

arises between the prosodic realization of information status and the referent's actual information status in discourse. Interestingly, a late positivity was observed only for the givenness accent. First, this indicates that de-accentuation – as a reflection of a highly active referent – is not adequate when the referent is available through a whole-to-part relation from an expression introduced in prior discourse, that is, when it has not yet been established as given information. Second, the absence of a positivity for the less expected newness accent (relative to the accessibility accent) implies that this accent type does not generate a conflict with the referent's information status: the referent represents new information in that it has not been mentioned in prior discourse, thus the newness accent is a permissible prosodic realization and does not impose a severe clash. This study thus provides evidence for a mismatch-induced positivity, showing that prosodically marked information status can conjure up a conflict between prosodic information and a referent's true information status. Converging with findings from reading comprehension, the data hence show that prosodic cues influence distinct stages of referential processing and that the language processor is sensitive to different prosodic realizations of information status.

Conclusion

This study reveals a biphasic N400-late positivity pattern during auditory referential processing. It shows that the N400 is sensitive to prosodic cues that support referent accessibility. The data critically indicate that the N400 is not merely influenced by semantic information but also by extra-semantic cues (i.e. accent types as markers of information status). The three-way N400 contrast shows that prosodic cues impact initial linking processes during the construction of a mental model. Therefore, the view that the N400 solely reflects lexico-semantic access cannot be maintained and should be broadened towards a more general account based on prominence relations that mark referent accessibility.

Acknowledgements

This research was supported by a grant from the German Research Foundation (DFG). We like to thank Anika Jödicke, Manuel Dangl, and Rainer Schmitz for their assistance during data preparation and acquisition.

References

1 Burkhardt P. Inferential bridging relations reveal distinct neural mechanisms: evidence from event-related brain potentials. *Brain Lang* 2006; **98**:159–168.
2 Burkhardt P. The P600 reflects cost of new information in discourse memory. *NeuroReport* 2007; **18**:1851–1854.
3 Chafe W. *Discourse, consciousness, and time*. Chicago/London: University of Chicago Press; 1994.
4 Burkhardt P, Roehm D. Differential effects of saliency: an event-related brain potential study. *Neurosci Lett* 2007; **413**:115–120.
5 Heim S, Alter K. Prosodic pitch accents in language comprehension and production: ERP data and acoustic analyses. *Acta Neurobiol Exp* 2006; **66**:55–68.
6 Magne C, Astesano C, Lacheret-Dujour A, Morel M, Alter K, Besson M. On-line processing of Pop-Out words in spoken French dialogues. *J Cogn Neurosci* 2005; **17**:740–756.
7 Toepel U, Pannekamp A, Alter K. Catching the news: processing strategies in listening to dialogs as measured by ERPs. *Behav Brain Funct* 2007; **3**:53.
8 Ladd DR. *Intonational phonology*. Cambridge: Cambridge University Press; 1996.
9 Cruttenden A. The de-accenting of given information: a cognitive universal? In: Bernini G, Schwartz ML, editors. *Pragmatic organization of discourse in the languages of Europe*. The Hague: Mouton de Gruyter; 2006. pp. 311–355.
10 Pierrehumbert JB, Hirschberg J. The meaning of intonational contours in the interpretation of discourse. In: Cohen PR, Morgan JL, Pollack ME, editors. *Intentions in communication*. Cambridge: MIT Press; 1990. pp. 271–311.
11 Kohler K. Terminal intonation patterns in single-accent utterances of German: phonetics, phonology and semantics. *AIPUK* 1991; **25**:115–185.
12 Baumann S, Grice M. The intonation of accessibility. *J Pragmat* 2006; **38**:1636–1657.
13 Röhr C, Baumann S. Prosodic marking of information status in German. *Speech Prosody* 2010. (in press).
14 Grice M, Baumann S, Benzmüller R. German intonation in autosegmental-metric phonology. In: Jun SA, editor. *Prosodic typology: the phonology of intonation and phrasing*. Oxford: Oxford University Press; 2005. pp. 55–83.
15 Boersma P, Weenink D. PRAAT, a system for doing phonetics by computer (version 3.4.). Amsterdam: Institute of Phonetic Sciences of the University of Amsterdam; 1996. *http://www.fon.hum.uva.nl/praat/*
16 Huynh H, Feldt LS. Conditions under which mean square ratios repeated measurements designs have exact F distributions. *J Am Stat Assoc* 1970; **65**:1582–1589.
17 Kutas M, Federmeier KD. Electrophysiology reveals semantic memory use in language comprehension. *Trends Cogn Sci* 2000; **4**:463–470.
18 Niedeggen M, Rösler F, Jost K. Processing of incongruous mental calculation problems: evidence for an arithmetic N400 effect. *Psychophysiology* 1999; **36**:307–324.
19 Sitnikova T, Kuperberg G, Holcomb PJ. Semantic integration in videos of real-world events: an electrophysiological investigation. *Psychophysiology* 2003; **40**:160–164.
20 Toepel U, Pannekamp A, van der Meer E. Fishing for information: the interpretation of focus in dialogs. In: Alter K, Horne M, Lindgren M, Roll M, Torkildsen JVK, editors. *Papers from Brain Talk*. Lund: Lund University; 2009. pp. 1–22.