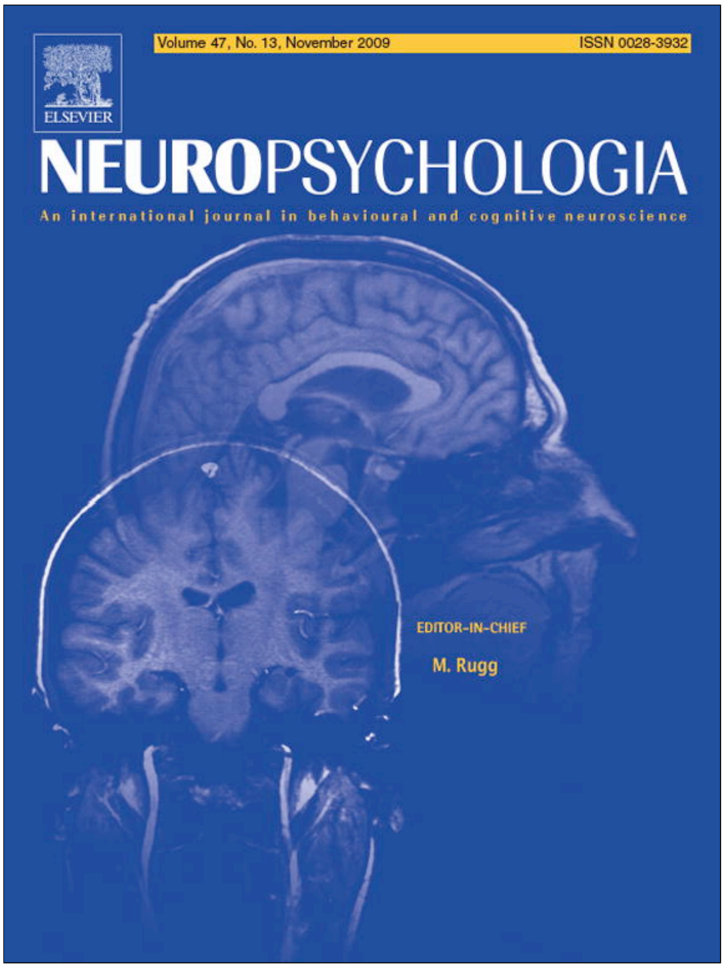
Volume 47, No. 13, November 2009
ISSN 0028-3932
ELSEVIER
NEUROPSYCHOLOGIA
An international journal in behavioural and cognitive neuroscience
EDITOR-IN-CHIEF
M. Rugg

Neuropsychologia 47 (2009) 3012–3022

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Note

The N400 as a correlate of interpretively relevant linguistic rules: Evidence from Hindi

Kamal Kumar Choudhary[a], Matthias Schlesewsky[b], Dietmar Roehm[a,c], Ina Bornkessel-Schlesewsky[a,d,*]

[a] Independent Junior Research Group Neurotypology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
[b] Department of English and Linguistics, Johannes Gutenberg-University, Mainz, Germany
[c] Department of Linguistics, University of Salzburg, Austria
[d] Department of Germanic Linguistics, University of Marburg, Germany

ARTICLE INFO

Article history:
Received 5 September 2008
Received in revised form 17 April 2009
Accepted 14 May 2009
Available online 22 May 2009

Keywords:
Language
N400
P600
LAN
Rules
Case marking
Ergative
Hindi

ABSTRACT

Classical views on the electrophysiology of language assume that different event-related potential (ERP) components index distinct linguistic subdomains. Hence, left-anterior negativities are often viewed as correlates of rule-based linguistic knowledge, whereas centro-parietal negativities (N400s) are taken to reflect (non-rule-based) semantic memory or aspects of lexical–semantic predictability. The present ERP study of case marking in Hindi challenges this clear-cut dichotomy. Though determined by a grammatical rule, the choice of subject case in Hindi is also interpretively relevant as it constrains the range of possible interpretations of the subject. For incorrect subject cases, we observed an N400, which was followed by a late positivity under certain circumstances. This finding suggests that violations of rule-based knowledge may engender an N400 when the rule is interpretively relevant.

1. Introduction

Since Kutas and Hillyard's (1980) seminal discovery of the first language-related event-related brain potential (ERP) component, the notion that different subdomains of linguistic knowledge can be associated with distinct ERP signatures has been a major driving force behind many electrophysiological investigations of human language processing. For example, ERP distinctions have featured prominently in the debate on whether rule-based linguistic knowledge can be dissociated from non-rule-based linguistic knowledge (McClelland & Patterson, 2002; Pinker & Ullman, 2002). In this context, transient left-anterior negativities (LANs) are often taken to index rule violations (cf. Ullman, 2004) whereas centro-parietal negativities within the same time range (N400 effects) are typically viewed as correlates of non-rule-based, lexically stored information (cf. Kutas & Federmeier, 2000). As an illustration, consider an ERP study by Weyerts, Penke, Dohrm, Clahsen, and Münte (1997), who observed LAN effects when a regular (default) plural suffix (*-s*) was illegally combined with an irregular noun stem in German (e.g. for **Bärs* vs. *Bären*, 'bears'), whereas the combination of a regular stem with an irregular plural suffix (*-en*) yielded an N400 (e.g. for **Wracken* vs. *Wracks*, 'wrecks'). Along with a number of similar findings in different languages (see Bornkessel-Schlesewsky & Schlesewsky, in press, for a recent overview), these results have been interpreted as showing that the overapplication of a morphological rule – as in "regularized" irregular words – correlates with a LAN, whereas the N400 in "irregularized" regulars reflects the fact that these words are treated like pseudowords (i.e. as they are non-decomposable, they result in a failure of lexical access because they do not have an entry in the mental lexicon).

A related component distinction is often drawn for language comprehension at the sentence level, with lexical–semantic integration difficulty or the degree of lexical–semantic predictability thought to be reflected in the N400, while other levels of linguistic analysis (e.g. syntax) are assumed to manifest themselves in qualitatively distinct ERP components such as LANs or late positivities, P600s (for a recent overview, see Kutas, Van Petten, & Kluender, 2006).

However, these relatively clear-cut functional interpretations of language-related ERP effects are challenged by several observations. First, LAN effects are not always observable in the face of violations that should be clearly rule-based (e.g. subject–verb

* Corresponding author at: Department of Germanic Linguistics, University of Marburg, Wilhelm-Roepke-Strasse 6A, 35032 Marburg, Germany.
Tel.: +49 06421 28 24 675; fax: +49 06421 28 24 558.
E-mail address: iboke@staff.uni-marburg.de (I. Bornkessel-Schlesewsky).

0028-3932/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neuropsychologia.2009.05.009

agreement, cf. Nevins, Dillon, Malhotra, & Phillips, 2007). Second, a number of relatively recent studies have reported P600 effects for seemingly *semantic* rather than syntactic manipulations (e.g. Hoeks, Stowe, & Doedens, 2004; Kim & Osterhout, 2005; Kolk, Chwilla, van Herten, & Oor, 2003; Kuperberg, Sitnikova, Caplan, & Holcomb, 2003). In view of findings such as these, some researchers favor more general interpretations of language-related ERP components, e.g. in terms of working memory-related processes for the LAN (e.g. King & Kutas, 1995; Kluender & Kutas, 1993) and of conflict monitoring for the P600 (e.g. van Herten, Chwilla, & Kolk, 2006; Vissers, Kolk, van de Meerendonk, & Chwilla, 2008).

For the N400, however, the non-rule-based, lexical–semantic perspective has been maintained more strongly. In particular, while it is clear than this component is modulated by a range of influencing factors (e.g. grammatical gender, Wicha, Moreno, & Kutas, 2004; or the broader discourse environment, cf. Hagoort & van Berkum, 2007, for an overview), all of these findings remain compatible with the notion that the N400 reflects the degree of predictability of a particular lexical item and the information associated with it (e.g. gender).[1] Notably, this view of the N400 is shared by a number of otherwise divergent neurocognitive models of language processing (e.g. Hagoort, 2005; Friederici, 2002; Ullman, 2004).

Yet there have also been reports of N400 effects elicited by manipulations that are not straightforwardly lexical–semantic in nature. For example, Frisch and Schlesewsky (2001) observed an N400 for a violation of case marking in German, and Bornkessel, McElree, Schlesewsky, and Friederici (2004) reported an N400 for word order reanalysis in certain types of German sentences (cf. also Haupt, Schlesewsky, Roehm, Friederici, & Bornkessel-Schlesewsky, 2008, for a demonstration of the reliability and generalizability of this finding).[2]

How should such "unexpected" component observations (Kutas et al., 2006) be interpreted? One possibility is to view the presence of an N400 as evidence for the non-rule-based nature of the linguistic manipulation in question (cf. Ye, Zhan, & Zhou, 2007, who argue for a lexically stored grammatical *construction* due to the finding of an N400). However, this line of argumentation brings with it the danger of circularity: since the classical functional interpretation of the N400 was based on a priori assumptions about which linguistic manipulations are rule-based/non-lexical or non-rule-based/lexical, it appears problematic to now apply the inverse line of argumentation to argue for the lexical status of manipulations that would traditionally be considered rule-based.

A second possibility is that LAN effects only index certain types of rule violations, namely those of *default* rules. For example, Bartke, Rösler, Streb, and Wiese (2005) observed N400 effects in response to the processing of morphological *subregularities* in German, i.e. to violations involving plural markers which are not the default form, but which are also not completely irregular. On the basis of this finding, these authors argued for an extension of the "rules vs. lexicon" distinction to a tripartite system in which rules are split up into default and non-default rules. The precise organization of such an architecture was left open, however.

Interestingly, there also appears to be a third possibility. When the experimental manipulations that have engendered classical, rule-based LAN effects are contrasted with those that have given rise to seemingly "rule-based" N400 effects, the difference between the two sets of studies could be characterized as follows: in those studies that elicited N400 effects, the critical rule violation was in some way relevant to sentence interpretation, whereas LAN effects appear to correlate with purely formal violations. For example, Frisch and Schlesewsky (2001) argued that the N400 response to the case violations in their study reflected a thematic interpretation problem such that the processing system could not determine "who is acting on whom". Likewise, the N400 effects observed for subject–object reanalyses (Bornkessel et al., 2004; Haupt et al., 2008) are clearly tied to interpretation at some level, because such reanalyses also involve a reinterpretation of which argument is the Actor (i.e. the participant primarily responsible for the state of affairs) and which is the Undergoer (i.e. the participant affected by the state of affairs). (For more details on the definition of Actor and Undergoer within a neurocognitive model of sentence comprehension, see Bornkessel & Schlesewsky, 2006; Bornkessel-Schlesewsky & Schlesewsky, 2009). By contrast, typical LAN-engendering manipulations like subject–verb agreement violations or overapplications of regular morphology are not associated with any clear interpretive consequences.

The aim of the present study was therefore to examine the hypothesis that seemingly "rule-based N400s" might be engendered by *interpretively relevant* rule-based information. To this end, we capitalized upon the properties of a language that has not yet been subjected to extensive neurocognitive investigation, namely Hindi.

2. The present study

As described in the preceding section, the classical functional interpretation of the N400 typically intertwines the notions of "non-rule-based" (or lexically stored) and "semantic", contrasting with rule-based, formal aspects of linguistic knowledge. This dichotomy between rules and linguistic form on the one hand and stored representations and meaning on the other has a high intuitive plausibility from the perspective of most European languages: in English, German or French, for example, violations of linguistic form (e.g. inflection, word order, case marking, agreement) do not straightforwardly appear to be associated with any consequences for word or sentence meaning. Yet, as we will describe in more detail below, this observation no longer appears so clear-cut when a broader range of languages is considered. For this reason, a recent neurocognitive model of language comprehension (the extended argument dependency model, eADM) posits that many of the neural correlates of language comprehension across typologically different languages are best captured via the assumption of a syntax–semantics *interface*, rather than in terms of a strict form vs. meaning dichotomy (Bornkessel & Schlesewsky, 2006; Bornkessel-Schlesewsky & Schlesewsky, 2008, 2009).

Hindi, an Indo-Aryan language spoken in India, is a case in point.[3] In this language, the case marking of subjects varies across sentence types: subjects are marked with nominative case in imperfective sentences and with the so-called ergative case in per-

[1] This does not necessarily mean that the N400 is a language-specific response: since graded N400 responses indexing the degree of semantic fit into a prior context have also been observed for non-verbal stimuli such a line drawings (e.g. Federmeier & Kutas, 2001; Nigam, Hoffman, & Simons, 1992), some authors interpret the N400 as a domain-general correlate of semantic memory use (Kutas & Federmeier, 2000). It should, however, be noted that N400 effects elicited by different types of stimuli (e.g. words vs. line drawings) are associated with somewhat different topographical distributions (Federmeier & Kutas, 2001). Hence, Kutas and Federmeier argue that this component "reflects the activity of a spatially distributed but temporally interlinked set of brain areas in both hemispheres [. . .], whose function is to bridge modality-specific sensory information and integrated, conceptual-level representations" (Kutas & Federmeier, 2000, p. 469).

[2] Note that the overall data pattern on subject–object reanalyses is somewhat more complex. For comprehensive overviews, see Bornkessel and Schlesewsky (2006) and Haupt et al. (2008).

[3] In 1995, an estimated number of 316 million people spoke Hindi or its sister language Urdu (spoken in Pakistan), thereby making it the language with the third-highest number of speakers globally (Graddol, 2004). Hindi is a verb-final language with a basic subject–object–verb (SOV) order, a relatively free word order (all six permutations of subject, object and verb are possible), and allowing argument drop.

Table 1
Examples for the four critical sentence types in the present study and per condition means for the two behavioral tasks (standard deviations are given in parentheses). Asterisks indicate that a sentence type violates the subject case marking rule. Condition labels encode the type of subject case (N = nominative; E = ergative) and the type of aspect (I = imperfective; P = perfective). ERPs were time-locked to the aspect marker *–taa*/*-tii* (3rd person singular masculine/feminine imperfective) or *–aa*/*ii* (3rd person singular masculine/feminine perfective), which is indicated in bold in the sentence examples and which provided the critical information as to whether the sentence involved a violation of subject case marking or not. Note that objects were always marked with the accusative marker *–ko*, which is required for all human direct objects in Hindi (e.g. Mohanan, 1994). While *–ko* is, in fact, ambiguous between accusative and dative case (i.e. it is also used to mark dative subjects and indirect objects of ditransitive verbs), this potential ambiguity appears unlikely to impact upon potential ERP differences between our critical conditions: the dative subject reading is ruled out by the presence of ergative case marking in conditions EI and EP and by the non-dative-initial word order in conditions NI and NP (cf. Lee, 2003, for a discussion of word order freezing with dative subjects in Hindi). If the possibility of an indirect object reading was not ruled out by prosodic phrasing, which is typically a good predictor for the upcoming clause–final verb in verb-final languages, this ambiguity between a direct and an indirect object reading will have been equal across conditions and therefore cannot account for any potential differences *between* our critical conditions. NOM: nominative; ACC: accusative; ERG: ergative; IPFV: imperfective; PFV: perfective; 3: 3rd person; M: masculine; AUX: auxiliary.

Critical sentence types					Acceptability judgment		Probe detection	
Condition	Example				Mean acceptability (%)	Mean reaction time (ms)	Mean accuracy (%)	Mean reaction time (ms)
NI	shikshak teacher.NOM "The teacher sees the gardener."	maalii-ko gardener-ACC	dekh-**taa** see-IPFV-3SG.M	hai AUX	93 (5)	480 (104)	96 (4)	822 (118)
EI	*shikshak-ne teacher-ERG "The teacher sees the gardener."	maalii-ko gardener-ACC	dekh-**taa** see-IPFV-3SG.M	hai AUX	8 (12)	469 (170)	95 (5)	843 (129)
NP	*shikshak teacher.NOM "The teacher has seen the gardener."	maalii-ko gardener-ACC	dekh-**aa** see-PFV-3SG.M	hai AUX	31 (24)	545 (177)	95 (5)	863 (131)
EP	shikshak-ne teacher-ERG "The teacher has seen the gardener."	maalii-ko gardener-ACC	dekh-**aa** see-PFV-3SG.M	hai AUX	92 (5)	471 (128)	95 (4)	814 (111)

fective sentences. This state of affairs is shown in Table 1, which also illustrates the critical sentence conditions for the present study.

As is apparent from Table 1, the case marking borne by the subject in Hindi is determined by the form of the verb. Like subject–verb agreement in English (cf. *The boy* <u>*runs*</u> vs. *The boys* <u>*run*</u>), this variation is typically described via a grammatical rule (e.g. Mohanan, 1994). In contrast to agreement, however, case marking is also relevant to the interpretation of the sentence participants (arguments). Whereas a nominative argument is compatible with a range of interpretations (including the Actor of a two-participant event as in NI in Table 1 or the only participant (Actor or Undergoer) in a one-participant event, cf. example 1), an ergative-marked argument must receive an agentive reading: as shown in Table 1, the ergative case marks the Actor of a two-participant event in the perfective aspect (i.e. Actor of a (completed) action, cf. EP). In a comprehensive theoretical examination of the Hindi/Urdu case system, Butt and King (2005) analyze this state of affairs by assuming that subjects are assigned nominative case by default, whereas the ergative is a "semantic" case that is assigned under certain, more restricted (but predictable) circumstances.

(1) शिक्षक बिमार हैं
sikshak bimaar hain
teacher ill is
'The teacher is ill.'

In summary, though the choice of subject case in Hindi is clearly governed by a grammatical rule, it is also associated with interpretive consequences because the range of possible interpretations of the ergative and the nominative case differ substantially. In accordance with this distinction, the choice of nominative vs. ergative subject case has been modeled via the assumption of a default vs. a non-default rule (Butt & King, 2005).

In the present study, we capitalized upon the interpretively relevant nature of the subject case marking rule(s) in Hindi in order to shed further light on the functional nature of the LAN/N400 distinction. Specifically, we examined the electrophysiological correlates of subject case marking violations by analyzing ERP responses time-locked to the position of the clause–final aspect marker of our critical sentences (marked in bold in the examples in Table 1). The rationale for time-locking directly to the aspect marker was that this should provide us with the most accurate approximation of the point in time at which the critical aspect information became available within the speech stream (for a similar line of reasoning, see van den Brink & Hagoort, 2004; Wolff, Schlesewsky, Hirotani, & Bornkessel-Schlesewsky, 2008).

As shown in Table 1, the experimental design crossed the factors case marking of the subject (CASE: ergative vs. nominative) and aspect (ASP: perfective vs. imperfective), thus giving rise to four critical conditions. The design therefore introduced violations in the two conditions in which the case marking and the aspect marking were not compatible (EI/NP). By manipulating whether the violation was effected by imperfective aspect marking following an ergative or by perfective aspect marking following a nominative, this design further allowed us to examine whether the strength or the nature of the violation response would be modulated by the misapplication of a default (nominative) vs. a non-default (ergative) rule.

Our hypotheses for the position of the aspect marker were as follows. If the violation of rule-based linguistic knowledge generally engenders LAN effects, our two violation conditions (NP/EI) should engender LANs in comparison to their grammatical controls (NI/EP). By contrast, if the nature of the rule-based knowledge is also important in determining the electrophysiological processing correlate, the interpretively relevant rule violations examined here might engender N400 effects. Thirdly, if the ERP correlate of a rule violation is conditioned by the type of linguistic rule (default vs. non-default, cf. Bartke et al., 2005), the overapplication of the nominative (default) rule should be expected to elicit a LAN in condition NP, whereas the overapplication of the ergative (non-default) rule in condition EI should be reflected in an N400. Finally, we expected to observe late positivities (P600 effects) for both violation conditions, as these typically occur in response to ill-formed sentences (Hagoort, Brown, & Groothusen, 1993; Osterhout & Holcomb, 1992).

3. Methods

3.1. Participants

Twenty-four native speakers of Hindi participated in the experiment after giving informed consent (7 women; mean age 26 years, range 19–38). All were right-handed, had normal or corrected-to-normal vision and good auditory acuity.

3.2. Materials

The sentence materials comprised 80 sets of the 4 sentence conditions in Table 1. Arguments were realized by human common nouns of the same gender (masculine in 40 sets and feminine in 40 sets). The 320 sentences thus resulting were subdivided into two lists of 160 sentences, each containing 40 sentences per condition and 2 sentences from a single lexical set. Each list was combined with 240 additional filler sentences and presented in a pseudo-randomized order, with list assignment counterbalanced across participants.

Sentences were spoken by a female native speaker of Hindi and recorded digitally (sampling rate: 44.1 kHz, 16 bit resolution). Acoustic analyses of the materials (see Appendix A for details) ensured that the violations introduced by the aspect markers in conditions EI and NP were not reflected in prosodic parameters prior to the onset of the verb. We subjected the following parameters for NP1, NP2, verb and auxiliary in each critical sentence to an item-based analysis of variance: duration, intensity and fundamental frequency (F_0) for constituent onset and offset and for the F_0-maximum and -minimum. Crucially, the acoustic analyses revealed no interaction of CASE and ASP prior to the critical verb position.

3.3. Procedure

Sentences were presented via loudspeakers. Each trial began with the presentation of a fixation asterisk (500 ms) before the onset of the sound file. After sentence offset, the asterisk remained on the screen for 1000 ms. Following a subsequent 500 ms of blank screen, participants completed two behavioral tasks using two hand-held push-buttons. Firstly, they judged whether the sentence that they had just heard was an acceptable sentence of Hindi or not. As a cue for the judgment task, three question marks appeared in the center of the computer screen. After a participant's response or after the maximal response time of 2500 ms had expired, a probe word appeared in the center of the screen. Participants were then required to judge whether this word had occurred in the previous sentence or not. The probe detection task required the answers 'yes' and 'no' equally often, with 'no' responses required in the case of an exchanged content word. The maximal response time for the probe detection task was also 2500 ms. After both tasks had been completed, there was a 2000 ms pause before the beginning of the next trial. The assignment of the left and right buttons to "yes" and "no" responses was counterbalanced across participants. In order to avoid ocular artifacts, participants were asked not to blink throughout the presentation of the fixation asterisk.

The experimental session was subdivided into 10 blocks of 40 sentences, between which participants took short breaks. A complete session lasted approximately 3 h including electrode preparation.

3.4. EEG recording and analysis

The EEG was recorded by means of 25 AgAgCl-electrodes (ground: AFZ). Recordings were referenced to the left mastoid (rereferenced to linked mastoids offline). The electro-oculogram (EOG) was monitored via electrodes at the outer canthus of each eye (horizontal EOG) and above and below the participant's right eye (vertical EOG). Electrode impedances were kept below 5 kΩ. EEG and EOG channels were amplified using a Twente Medical Systems DC amplifier (Enschede, The Netherlands) and recorded with a digitization rate of 250 Hz.

In order to eliminate slow signal drifts, a 0.3–20 Hz band-pass filter was applied to the raw EEG data. Subsequently, average ERPs were calculated per condition per participant from the onset of the aspect marker to 1000 ms post onset, before grand averages were computed over all participants. No baseline corrections were performed, in order to avoid a distortion of the critical ERP epochs via possible transient signal differences in the baseline interval. This issue arises particularly when employing auditory presentation and acoustically unaltered auditory materials, because here the stimulus materials by definition cannot be fully identical before critical word onset (for a similar method of data analysis based on the same rationale, see Friederici, Wang, Herrmann, Maess, & Oertel, 2000; Haupt et al., 2008; Philipp, Bornkessel-Schlesewsky, Bisang, & Schlesewsky, 2008; Wolff et al., 2008).

Trials for which the probe detection task was not performed correctly were excluded from the averaging procedure, as were trials containing EEG or EOG artifacts (the EOG rejection criterion was 40 μV). The number of trials entering the averaging procedure did not differ across conditions (mean/standard deviation per condition: EI = 29/7; EP = 31/6; NI = 31/7; NP = 30/6; $F < 1$).[4]

For the statistical analysis of the ERP data, repeated-measures ANOVAs were calculated for mean amplitude values per time window per condition in four regions of interest (ROIs). Lateral ROIs were defined as follows: *left-anterior* (F3, F7, FC1, FC5); *left-posterior* (CP1, CP5, P3, P7); *right-anterior* (F4, F8, FC2, FC6); *right-posterior* (CP2, CP6, P4, P8). For the midline electrodes, each electrode (FZ, FCZ, CZ, CPZ, PZ, POZ) was treated as a ROI of its own. Time windows were chosen on the basis of visual inspection; since no previous auditory ERP studies of Hindi have been reported, it was virtually impossible to employ pre-defined time windows on the basis of previous studies. For analyses involving more than one degree of freedom in the numerator, significance values were corrected when sphericity was violated (Huynh & Feldt, 1970).

4. Results

4.1. Behavioral data

Mean acceptability ratings/error rates and reaction times for the two behavioral tasks are shown in the center and right-hand panels of Table 1.

For the acceptability task, we chose to conduct the statistical analysis on mean deviations from the expected judgment (i.e. deviations from "acceptable" for EP/NI and from "unacceptable" for EI/NP) rather than on mean acceptability ratings in order to better capture the rating difference between the two ungrammatical conditions. In this regard, repeated-measures ANOVA (by participants (F_1) and items (F_2)) revealed main effects of ASP ($F_1(1,23) = 26.01$, $p < 0.001$; $F_2(1,79) = 88.93$, $p < 0.001$) and CASE ($F_1(1,23) = 22.77$, $p < 0.001$; $F_2(1,79) = 63.81$, $p < 0.001$) as well as an interaction of ASP and CASE ($F_1(1,23) = 8.78$, $p < 0.007$; $F_1(1,79) = 44.74$, $p < 0.001$). Resolving the interaction by ASP showed simple effects of CASE for the perfective ($F_1(1,23) = 15.56$, $p < 0.0006$; $F_2(1,79) = 67.64$, $p < 0.001$) but not for the imperfective sentences.

The analysis of the error rates for the probe detection task did not yield any significant effects.

We refrain from analyzing the reaction times as the tasks were not directly time-locked to the critical region of the sentence.

4.2. ERP data

4.2.1. ERPs time-locked to the critical aspect marker

Grand average ERPs time-locked to the onset of the critical aspect marker are shown in Fig. 1. As is apparent from the figure, the conditions involving violations of subject case marking (EI/NP) engendered centro-parietal negativities between approximately 100 and 300 ms post onset of the aspect marker. In condition EP, this negativity was followed by a large positivity with a centro-parietal maximum (approximately 400–700 ms).

In the 100–300 ms window, a repeated-measures ANOVA revealed a main effect of aspect for the midline electrodes ($F(1,23) = 4.38$, $p < 0.04$) and interactions of ROI*CASE (midline: $F(5,115) = 6.46$, $p < 0.002$; lateral: $F(3,69) = 5.71$, $p < 0.01$), CASE*ASP (midline: $F(1,23) = 14.47$, $p < 0.001$; lateral: $F(1,23) = 12.12$, $p < 0.01$) and ROI*CASE*ASP (midline: $F(5,115) = 7.57$, $p < 0.001$; lateral: $F(3,69) = 7.57$, $p < 0.001$). Resolving these interactions by ROI revealed interactions of CASE*ASP at electrodes FCZ ($F(1,23) = 6.40$, $p < 0.01$), CZ ($F(1,23) = 21.14$, $p < 0.001$), CPZ ($F(1,23) = 19.09$, $p < 0.001$), PZ($F(1,23) = 16.69$, $p < 0.001$) and POZ ($F(1,23) = 18.42$, $p < 0.001$) and in the left-posterior ($F(1,23) = 17.08$, $p < 0.001$), right-posterior ($F(1,23) = 23.74$, $p < 0.001$) and right-anterior ($F(1,23) = 5.27$, $p < 0.05$) regions. With the exception of the midline electrode CPZ, all of the ROIs showing an interaction CASE*ASP also showed an effect of CASE for *imperfective* sentences (CZ: $F(1,23) = 16.10$, $p < 0.001$; CPZ: $F(1,23) = 15.47$, $p < 0.001$;

[4] For the additional analysis relative to the onset of the verb which will be reported later, the average trial numbers/standard deviations per condition were as follows: EI = 28/7; EP = 31/6; NI = 30/7; NP = 30/7. Again, there was no significant difference between conditions ($F < 1$).

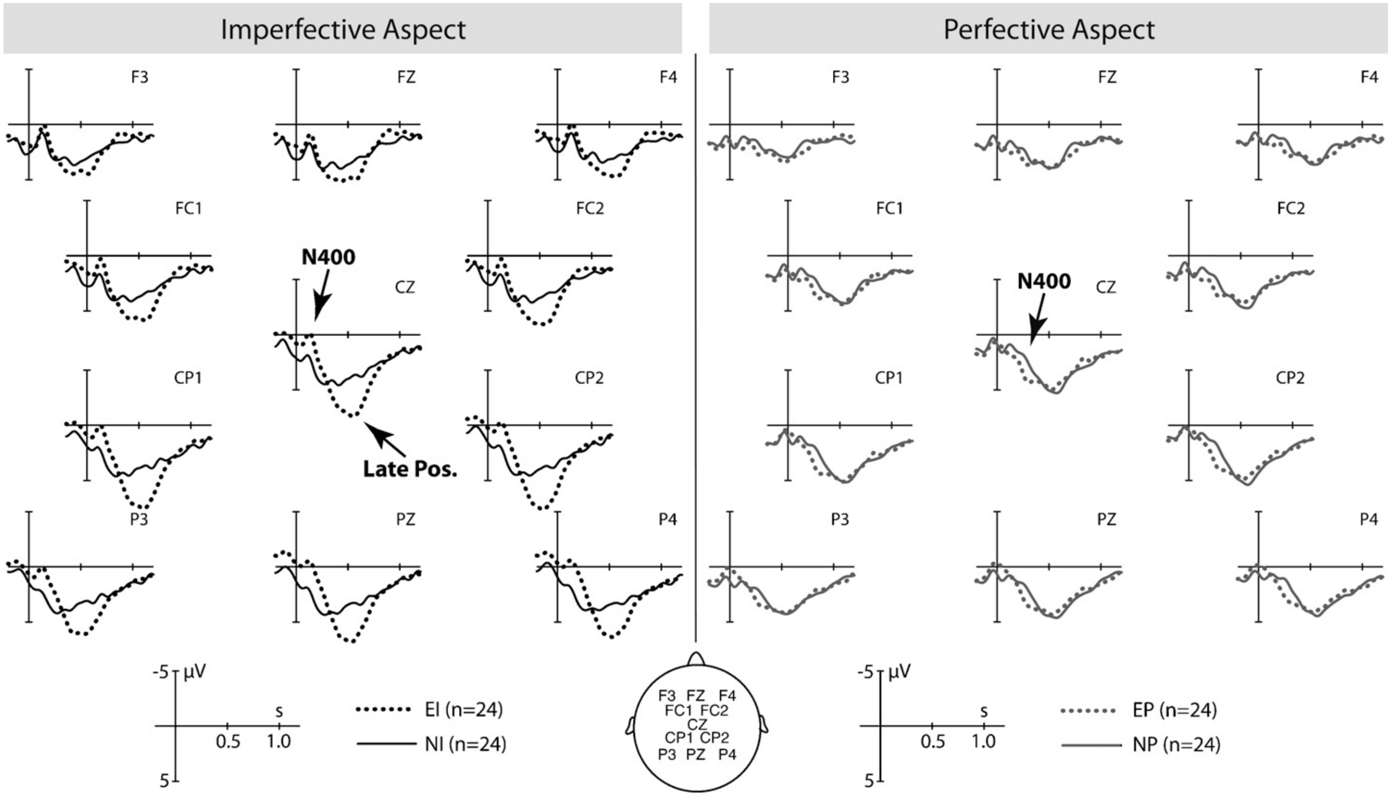

Fig. 1. Grand average ERPs ($N = 24$) time-locked to the onset of the critical aspect marker (onset at the vertical bar). The left panel shows comparisons between imperfective sentences with correct (NI) and incorrect (EI) subject case marking. The right panel shows comparisons between perfective sentences with correct (EP) and incorrect (NP) subject case marking. Negativity is plotted upwards.

PZ: $F(1,23) = 18.09$, $p < 0.001$; POZ: $F(1,23) = 16.69$, $p < 0.001$; left-posterior: $F(1,23) = 22.69$, $p < 0.001$; right-posterior: $F(1,23) = 26.33$, $p < 0.001$; right-anterior: $F(1,23) = 4.03$, $p < 0.05$). In all cases, the effect was due to a negativity for sentences with an ergative first NP. For perfective sentences, by contrast, the effect of CASE only reached significance at electrodes FCZ ($F(1,23) = 4.04$, $p < 0.001$), CZ ($F(1,23) = 9.07$, $p < 0.006$), CPZ ($F(1,23) = 9.96$, $p < 0.004$) and marginal significance in the right-posterior region ($F(1,23) = 3.73$, $p < 0.06$). Here, the effects were due to a negativity for sentences with a nominative first NP.

For the 400–700 ms time window, the statistical analysis revealed main effects of ASP (midline: $F(1,23) = 8.85$, $p < 0.001$; lateral: $F(1,23) = 11.92$, $p < 0.01$) and CASE (midline: $F(1,23) = 32.37$, $p < 0.001$; lateral: $F(1,23) = 20.84$, $p < 0.001$) as well as interactions of CASE*ASP (midline: $F(1,23) = 25.48$, $p < 0.001$; lateral: $F(1,23) = 15.90$, $p < 0.001$) and ROI*CASE*ASP (midline: $F(5,115) = 7.60$, $p < 0.002$; lateral: $F(3,69) = 9.24$, $p < 0.001$). Resolving these interactions by ROI revealed interactions of CASE*ASP at all midline sites (maximum at CPZ: $F(1,23) = 35.50$, $p < 0.001$; minimum at FZ: $F(1,23) = 4.67$, $p < 0.04$) and in the left-posterior, right-posterior and right-anterior regions (all $Fs(1,23) > 10$, all $ps < 0.01$). All of the ROIs showing this interaction also showed a significant effect of CASE for imperfective sentences (all $Fs(1,23) > 10$, all $ps < 0.004$). By contrast, no effects of CASE were observable for perfective sentences in this time window.

4.2.2. Additional analyses

4.2.2.1. Analysis of ERPs relative to verb onset. The ERP analyses relative to the aspect marker revealed a negativity with a classical N400 distribution (posterior maximum with a right-hemispheric focus, cf. Kutas & Federmeier, 2000). At a first glance, however, its latency (100–300 ms) appears too short for an N400. One possibility is that this temporal discrepancy is due to the fact that analyses were time-locked directly to the critical aspect marker and not to word onset (as is typically the case in the N400 literature). In order to examine this hypothesis, we computed ERPs relative to the onset of the verb. These are shown in Fig. 2 for electrode CZ. The figure also marks the mean onset of the aspect marker in perfective and imperfective sentences, thus allowing for a better comparability with Fig. 1.

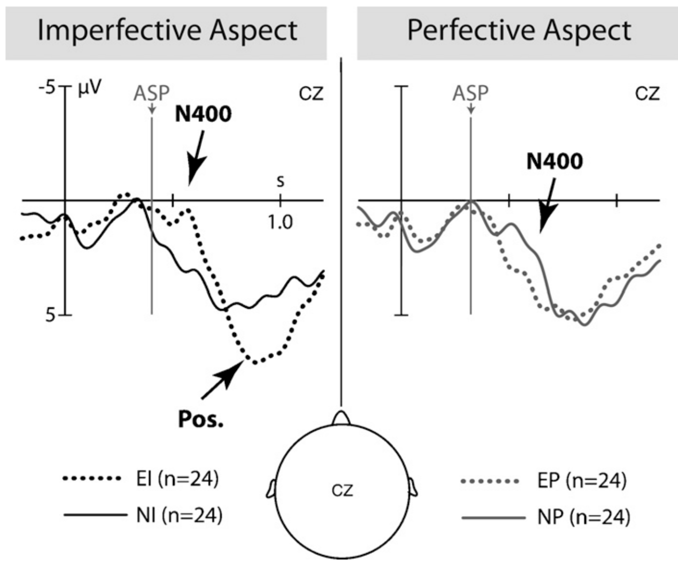

Fig. 2. Grand average ERPs ($N = 24$) time-locked to the onset of the verb (onset at the vertical bar) at one selected electrode. The left panel shows comparisons between imperfective sentences with correct (NI) and incorrect (EI) subject case marking. The right panel shows comparisons between perfective sentences with correct (EP) and incorrect (NP) subject case marking. The approximate mean onsets of the aspect marker are indicated by vertical grey lines. Negativity is plotted upwards.

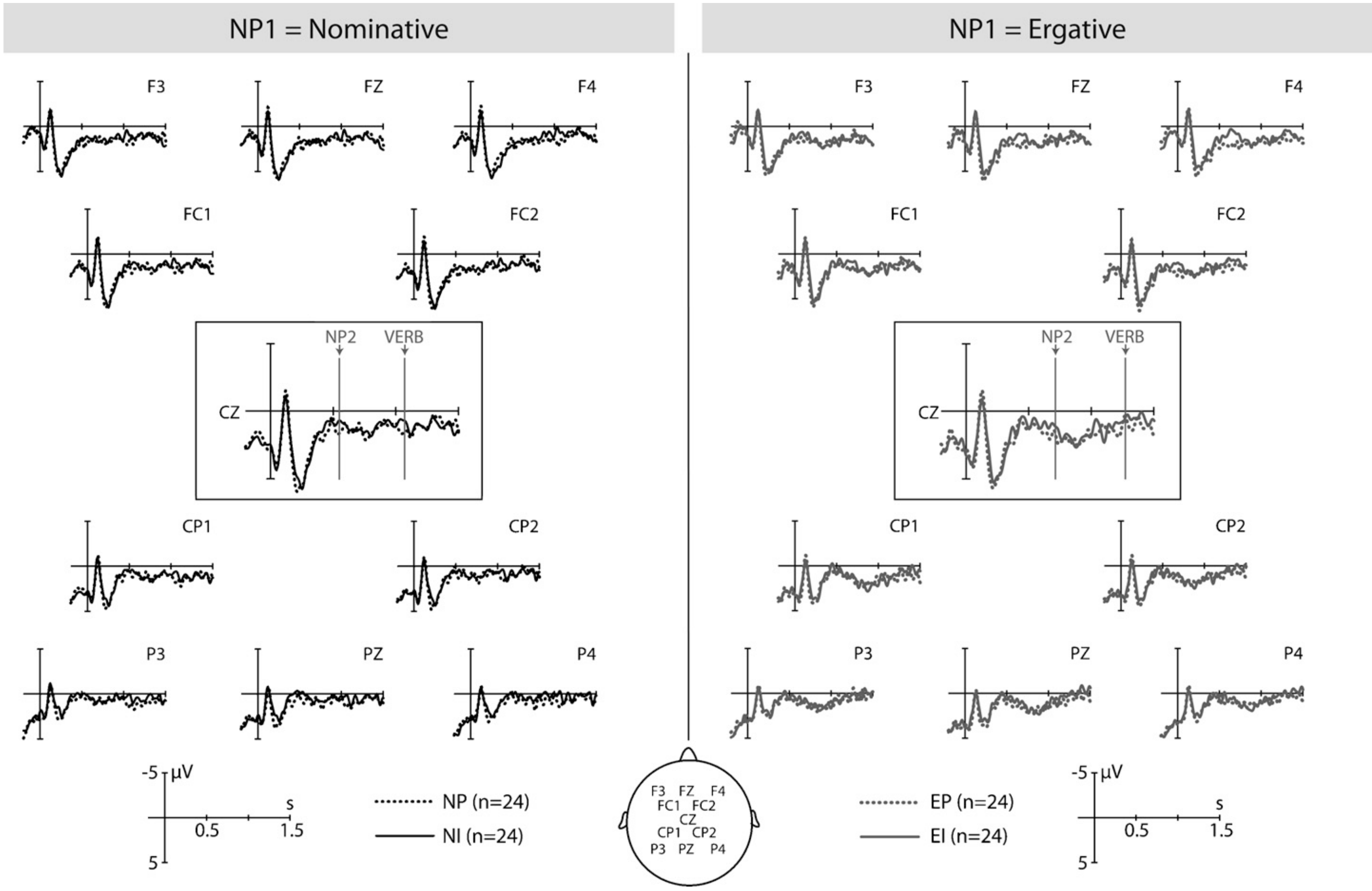

Fig. 3. Grand average ERPs ($N = 24$) time-locked to the onset of NP1 (onset at the vertical bar) and spanning both NP1 and NP2. ERPs are plotted separately for sentences with a nominative first NP (left panel) and for sentences with an ergative first NP (right panel) due to the differences in the duration of NP1 in the two cases (see Appendix A). The approximate mean onsets of NP2 and the verb are indicated by vertical grey lines. Negativity is plotted upwards.

As Fig. 2 shows, the ERP responses to subject case marking violations appear in much more typical time windows for N400 and P600 components when measured relative to verb onset. Statistical analyses in the time windows 450–700 and 750–1100 ms relative to verb onset (see Appendix B) confirmed the results of the analyses relative to the onset of the aspect marker. Hence, the analysis of ERPs relative to verb onset supports an N400-interpretation of the negativity.

4.2.2.2. Analysis of ERPs to the pre-critical regions (NP1, NP2). As our experimental sentences were presented as natural, unaltered speech, it is important to determine whether the ERP responses differed prior to the critical point in the sentence (i.e. prior to the verb/aspect marker). To this end, we computed ERPs relative to the onset of NP1 and spanning both NP1 and NP2 up to the average point of verb onset. Fig. 3 shows these ERP responses separately for sentences with a nominative and an ergative-marked NP1 due to the different durations of NP1 in both cases (see Appendix A). Onsets of NP2 and the verb are indicated in the figure.

Fig. 3 shows that ERP responses to ergative- and nominative-initial sentences were very similar prior to the verb; a descriptive impression that was confirmed by statistical analyses in consecutive time windows. (Note especially that the apparent right-anterior negativity for condition EI between approximate 500 and 700 ms post onset of NP1 did not reach significance. This effect was mainly visible at a single electrode, namely F4, and is thus likely an artifact.) The statistical analysis of the regions prior to the verb only revealed a single significant effect, namely a late positivity for ergative vs. nominative first arguments relative to the position of NP1. In order to show this effect more clearly, Fig. 4 depicts the ERPs engendered by sentences with an ergative and a nominative NP1, respectively, collapsing over aspect.

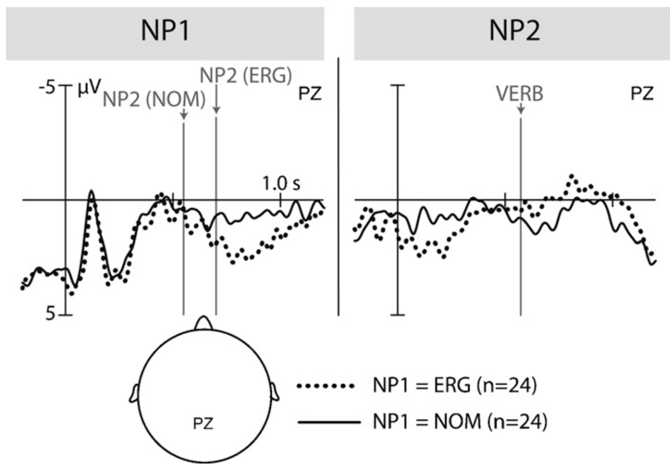

Fig. 4. Grand average ERPs ($N = 24$) time-locked to the onset of NP1 (left panel) and NP2 (right panel) for sentences with nominative and ergative first NPs at a single selected electrode. For the left panel (ERPs relative to NP1), the approximate mean onsets of NP2 are indicated by grey vertical lines for sentences with nominative and ergative first NPs, respectively. For the right panel (ERPs relative to NP2), the approximate mean onset of the verb is indicated by a grey vertical line. Negativity is plotted upwards.

As is apparent from Fig. 4, the positivity for initial ergative vs. nominative NPs indeed appears to stem from the processing of NP1 – rather than being an early effect of the processing of NP2 – since it begins a little before the onset of NP2. The statistical analysis in a time window from 650 to 1100 ms relative to the onset of NP1 revealed a main effect of CASE (lateral: $F(1,23) = 8.91$, $p < 0.01$; midline: $F(1,23) = 22.77$, $p < 0.001$), and, for the midline electrodes, an interaction ROI × CASE ($F(5,115) = 5.06$, $p < 0.02$). The simple effect of CASE reached significance at all electrodes from FCZ to POZ, with a maximum at POZ ($F(1,23) = 23.77$, $p < 0.001$). It is possible that this effect reflects the different acoustic properties of nominative and ergative first NPs (e.g. in terms of length, see Appendix A). Crucially, however, as it did not interact with aspect, it cannot account for our critical findings at the position of the clause–final verb/aspect marker.

5. Discussion

The present auditory ERP study of subject case marking violations in Hindi revealed a negativity with a classical N400 distribution (centro-parietal maximum) for both violation conditions. The effect was more pronounced for the violation involving ergative case combined with imperfective aspect, as was apparent from its broader distribution. Furthermore, in the ergative-imperfective condition, the negativity was followed by a broadly distributed positivity with a parietal maximum. Whereas the negativity and the positivity showed relatively short latencies relative to the onset of the aspect marker (100–300 ms for the negativity; 400–700 ms for the positivity), analyses relative to the onset of the verb revealed latency ranges (450–700 ms; 750–1100 ms) that are much more typical for N400 effects and late positivities (P600s), respectively. (Note that, while slight differences between conditions are observable prior to the onset of the aspect marker, these presumably resulted from intra-word coarticulation, on the basis of which the comprehension system may have been able to anticipate the aspect information prior to the onset of the aspect marker on some trials.) These findings are compatible with previous observations of "latency-shifted" language-related ERP components under auditory presentation conditions (cf. van den Brink & Hagoort, 2004, for an early N400 onset via phonological information; and Kutas et al., 2006, for a discussion of N400 onsets as early as 50 ms in natural speech). They are also highly comparable to recent findings from Japanese, in which ERP effects that appeared with a very short onset latency (approximately 100 ms) when ERPs were time-locked to a critical morpheme in auditory presentation were observed in considerably later time windows (from 300 ms onwards) when the identical sentence materials were presented visually (Wolff et al., 2008). All of these previous findings thus suggest that typical ERP component latencies would likely be considerably shorter if it were always possible to time-lock directly to the onset of the critical information. These considerations and the topographical distribution of the negativity thus justify a classification of the negativity effects in the present study as N400 effects. In the following, we will first discuss the functional significance of the N400 before turning to the late positivity.

5.1. N400 effects

The present study aimed to contrast three hypotheses with regard to the processing of rule-based linguistic knowledge and its electrophysiological correlates: (a) the traditional notion that rule violations generally engender LAN effects rather than N400s; (b) the modified perspective that the violation of default rules elicits LANs, whereas the violation of non-default rules may correlate with N400s; and (c) an alternative proposal according to which interpretively relevant rules may elicit N400 effects. Our findings clearly support the third hypothesis, as both types of subject case marking violations in Hindi engendered N400 effects. Thus an N400 was observed independently of whether the rule violation involved a default rule (nominative case assignment) or a non-default rule (ergative case assignment). Recall from the introduction that, in addition to being governed by the aspect of the sentence, the choice of ergative or nominative case changes the range of possible interpretations that may be assigned to the subject. Our findings indicate that under these interpretively relevant circumstances, rule violations correlate with N400s rather than LAN effects. This observation is not straightforwardly compatible with the notion that N400s correlate with the processing of lexically stored knowledge/semantic memory nor with the assumption that they reflect lexical–semantic predictions, since both of these perspectives make reference to activation changes of individual lexical items and the information tied to them (e.g. gender) as opposed to abstract grammatical information (e.g. the required aspect marking in the present study).

While the default vs. non-default status of the subject case marking rule did not affect the type of negativity observed, it did impact upon the magnitude of the N400 effect and, in addition, modulated the presence or absence of the following positivity. With regard to the N400, it is interesting to note that the default vs. non-default status of the case marking rule correlates with the number of interpretive possibilities that are associated with the case marker in question. In other words, the processing conflict reflected in the N400 is stronger for the case marker with the more constrained range of possible interpretations (ergative).[5] This observation thus supports the idea that the N400 effects in the present study correlate with the processing of interpretively relevant grammatical rules: the greater the interpretive problem that is engendered by the rule misapplication, the more pronounced the negativity. Indeed, a similar suggestion was put forward by Frisch and Schlesewsky (2001, 2005), who reported N400 effects in response to "double case violations" in German. These violations involve two identically case-marked arguments, thereby giving rise to a thematic interpretation conflict ("who is acting on whom?"). Frisch and Schlesewsky (2005) observed a reduction of N400 amplitude when the case marking that engendered the violation was compatible with a broader range of possible interpretations (nominative and dative vs. accusative in German). A similar observation holds when the two arguments differed with respect to an additional feature (animacy) that could help to resolve the interpretation conflict (Frisch & Schlesewsky, 2001). (For an alternative proposal, see de Hoop & Lamers, 2006; Lamers & de Hoop, 2005.) An account along these lines receives further support from the observation – as revealed by a comparison of ERPs with speed–accuracy trade-off measures – that N400 effects can reflect both conflict detection and conflict resolution (Bornkessel et al., 2004; Experiment 3).

In summary, the present findings provide strong converging support for the perspective that the processing of rule-based linguistic knowledge correlates with an N400 when the consequences of a rule misapplication are interpretive, rather than purely formal in nature. This suggests that the traditional functional dichotomy in the cognitive neuroscience of language, in which rule-based and syntactic/morphological information on the one hand are contrasted with lexical and semantic information on the other, needs to be reconsidered—or at least refined. When a broader range of rules in a wider range of languages is considered, it becomes apparent that rule-based knowledge need not always be without interpretive

[5] The more flexible interpretive possibilities for nominative case are also reflected in our participants' acceptability ratings: condition NP (with an incorrect nominative) was judged to be considerably more acceptable (31%) than condition EI (with an incorrect ergative; 8%).

consequences and that it is the distinction between interpretively relevant and interpretively irrelevant that appears responsible for engendering the dissociation between the N400 and other effects rather than the distinction between non-rule-based/lexical and rule-based knowledge.

5.2. The late positivity

In addition to N400 effects for both violation conditions, the present study also showed the predicted late positivity (P600) for the violation condition with an incorrect ergative (EI). Unexpectedly, however, the second violation condition (NP) did not show such an effect. This striking – and unpredicted – difference between the two violation conditions may again be related to differences between the two types of subject case markers. In addition to the N400 modulations described above, Frisch and Schlesewsky's (2005) results on the processing of double case violations in German revealed more pronounced late positivities for violations involving dative as opposed to nominative and accusative case. Crucially, since accusative is the default object case in German and dative is the exception, this observation is highly compatible with the present findings, in which the overapplication of the non-default (ergative) rule yielded a late positivity whereas the overapplication of the default (nominative) rule did not.

Yet, while differences in the restrictiveness of the two subject cases can account for the attenuation of the positivity in NP vs. EI, they do not explain why NP did not engender a positivity at all. There is, however, an additional fact about Hindi that might be relevant in this regard, namely that a small number of transitive verbs (e.g. *laanaa*, 'to bring') exceptionally require nominative subjects even in the perfective aspect (e.g. Butt & King, 2005; Mohanan, 1994). Though none of these verbs were used in the materials of the present study, their existence may be of importance for the behavior of the language comprehension system: while the occurrence of an ergative subject in a transitive, imperfective sentence is completely ruled out, a nominative subject may occur in a transitive, perfective sentence under certain (though very restricted) circumstances. From this perspective, the late positivity could reflect the *principled incompatibility* between the subject case and the aspect marker. Such an incompatibility only arises in condition EI but not in condition NP.

Though an interpretation along these lines is clearly speculative at present and will require further verification, the observation that grammatical violations need not necessarily engender a late positivity is problematic for most functional accounts of the P600/late positivity. This effect has been described as a marker of syntactic processing (Hagoort, 2003; Ullman, 2004), of reanalysis/repair (Friederici, 2002), of well-formedness problems (Bornkessel & Schlesewsky, 2006) or of more general conflict detection (van Herten et al., 2006). The absence of a late positivity in condition NP does not follow straightforwardly from any of these approaches, thereby suggesting that a more fine-grained functional differentiation may be required.[6] The tentative interpretation advanced above may help to provide a first step in this direction: it yields the testable hypothesis that grammatical exceptions may be just as important as grammatical rules in shaping the neurophysiological signatures of language comprehension.

6. Conclusion

The present findings indicate that, in contrast to classical assumptions about the electrophysiology of language, the processing of grammatical rules may be reflected in N400 effects. Hence, the distinction between LAN and N400 effects cannot be taken as an indication of rule-based vs. lexically stored linguistic knowledge. Rather, the *type* of rule-based knowledge may determine the electrophysiological processing signature. Whereas the purely formal rule violations that have most typically been examined in neurocognitive studies (e.g. subject–verb agreement) engender LAN effects, rule violations that touch upon interpretively relevant information (e.g. subject case marking in Hindi) may lead to N400 effects. Furthermore, our results provide a first indication that late positive ERP effects (P600s) may be highly sensitive to rule exceptions (i.e. occur only in response to principled incompatibilities between grammatical features such as case and aspect). These findings highlight the importance of examining non-European languages in order to test hypotheses about the neural language processing architecture.

Acknowledgements

The research reported here was supported by a grant from the German Research Foundation (BO 2471/2-1/SCHL 544/2-1) and was conducted in collaboration with the Clinic for Audiology and Phoniatry (Prof. Manfred Gross) of the Charité Berlin. We are grateful to Katja Bruening for invaluable assistance in data acquisition and to Kerstin Flake for expert assistance in the preparation of the figures.

Appendix A. Acoustic analyses

For the analyses, the following parameters were extracted for NP1, NP2, verb and auxiliary in each critical sentence: duration, intensity and fundamental frequency (F_0) for constituent onset and offset and for the F_0-maximum and -minimum. Mean values for duration and intensity are given in Table 2 and pitch contours are visualized in Fig. 5.

In order to examine whether the acoustic parameters of our critical stimuli differed across conditions, we conducted item-based ANOVAs. Crucially, the results of these analyses – which are reported in more detail in the following – revealed no interactions of CASE and ASP that could account for our critical ERP findings.

The analysis of the mean durations revealed a main effect of CASE for NP1 ($F(1,79) = 458.28$, $p < 0.001$), reflecting the increased length of ergative first arguments. For the pause between NP1 and NP2, the main effect of CASE reached marginal significance ($F(1,79) = 3.92$, $p < 0.06$) due to a slightly longer pause for the shorter, nominative first arguments. At the position of the verb, the analysis showed a main effect of ASP ($F(1,79) = 111.72$, $p < 0.001$), which was due to the longer imperfective markers. In addition, the interaction of CASE*ASP reached marginal significance at the verb position ($F(1,79) = 3.67$, $p < 0.06$), but resolving the interaction by ASP did not reveal a significant effect of CASE for either level of aspect. The main effect of ASP was also observable for the pause between the verb and the auxiliary ($F(1,79) = 7.25$, $p < 0.01$) and for the auxiliary ($F(1,79) = 9.71$, $p < 0.01$). In both cases, however, the imperfective aspect showed shorter durations.

[6] One could, however, envisage variants of the well-formedness and conflict monitoring approaches that might be able to account for these findings. Notably, what counts as a "well-formed" or "conflict free" utterance is highly dependent on the communicative environment. This may play a substantial role with respect to Hindi, which is the official language of Northern India and is therefore used for communication by speakers with many different native languages. Since not all of these languages have ergative case marking, the omission of ergative case in sentences with perfective aspect is not uncommon – as pointed out by our informants. By contrast, the overapplication of ergative case marking to non-perfective sentences does not occur. This plurality of dialectal variants – perhaps in concert with the notion of principled (in)compatibility introduced above – could conceivably lead to a modulation within the well-formedness evaluation, which was conceived as a task- and environment-dependent process (Bornkessel and Schlesewsky, 2006).

Table 2
Mean durations and intensities for each of the constituents in the four critical sentence conditions. Standard deviations are given in parentheses.

Condition	Mean intensity (dB)				Mean duration, constituents (ms)				Mean duration, pauses (ms)		
	NP1	NP2	Verb	Aux	NP1	NP2	Verb	Aux	NP1–NP2	NP2–Verb	Verb–Aux
NI	64.5 (3.5)	64.4 (3.6)	59.2 (2.4)	57.7 (2.2)	516 (108)	554 (83)	565 (60)	180 (21)	35 (47)	20 (20)	12 (10)
EI	64.1 (3.5)	63.6 (3.8)	58.0 (2.6)	56.1 (2.1)	677 (107)	550 (83)	551 (87)	179 (23)	29 (33)	30 (66)	11 (7)
NP	64.3 (3.6)	64.4 (3.7)	63.1 (2.4)	57.5 (1.9)	509 (116)	555 (93)	491 (71)	189 (25)	42 (52)	28 (23)	15 (15)
EP	63.8 (3.4)	62.9 (3.9)	61.2 (2.6)	56.0 (1.9)	669 (97)	557 (82)	495 (84)	187 (27)	33 (35)	24 (22)	16 (12)

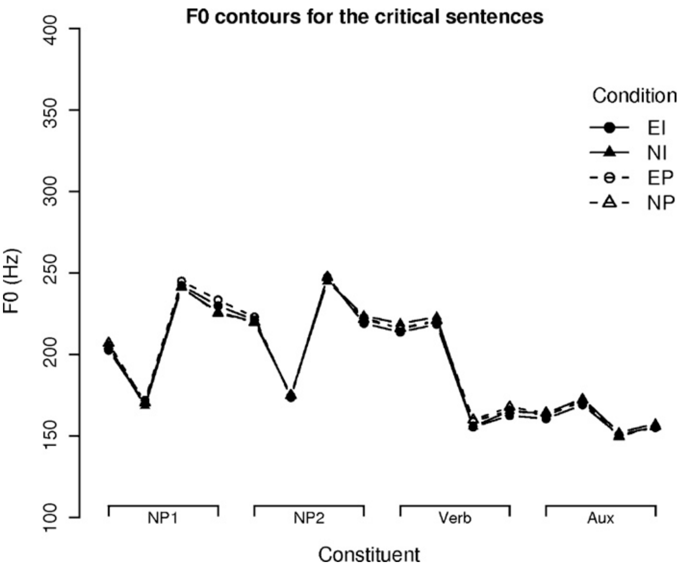

Fig. 5. F_0 contours for the four critical conditions.

The analysis of the mean intensities showed a main effect of CASE for NP1 ($F(1,79) = 34.36$, $p < 0.001$) and for NP2 ($F(1,79) = 82.72$, $p < 0.001$). For the verb, there were main effects of ASP ($F(1,79) = 465.61$, $p < 0.001$), CASE ($F(1,79) = 60.18$, $p < 0.001$) and an interaction CASE*ASP ($F(1,79) = 5.36$, $p < 0.03$). Resolving the interaction by ASP showed main effects of CASE for both the imperfective ($F(1,79) = 24.68$, $p < 0.001$) and the perfective ($F(1,79) = 53.54$, $p < 0.001$) aspect. Crucially, however, the interaction was not due to a reversal of the case effect between the two levels of the factor aspect, but rather to a larger effect in the perfective aspect. Hence, this interaction cannot account for the differential ERP effects at the position of the verb/aspect marker. Finally, the analysis of the auxiliary showed a main effect of CASE ($F(1,79) = 56.16$, $p < 0.001$). Note also that all intensity differences were exceedingly small (<3 dB).

The analysis of the F_0 contours did not reveal any effects that exceeded the threshold for perception (Rietveld & Gussenhoven, 1985; t'Hart, Collier, & Cohen, 1990).

Appendix B. ERP statistics relative to the onset of the verb

B.1. 450–700 ms

The analysis for the 450–700 ms time window showed an interaction of ROI*CASE, which only reached significance for the midline electrodes ($F(3,69) = 3.36$, $p < 0.03$). Both midline and lateral sites showed interactions of CASE*ASP (midline: $F(1,23) = 16.78$, $p < 0.001$; lateral: $F(1,23) = 15.49$, $p < 0.001$) and ROI*CASE*ASP (midline: $F(5,115) = 5.47$, $p < 0.01$; lateral: $F(3,69) = 9.76$, $p < 0.001$). Resolving the interactions by ROI showed an interaction of CASE*ASP for the midline electrodes FCZ ($F(1,23) = 8.18$, $p < 0.008$), CZ ($F(1,23) = 24.51$, $p < 0.001$), CPZ ($F(1,23) = 21.17$, $p < 0.001$), PZ ($F(1,23) = 16.14$, $p < 0.001$) and POZ ($F(1,23) = 13.31$, $p < 0.001$) and for the left-posterior ($F(1,23) = 23.76$, $p < 0.001$), right-posterior ($F(1,23) = 23.28$, $p < 0.001$) and right-anterior ($F(1,23) = 5.83$, $p < 0.02$) regions.

The interactions in individual ROIs were further resolved by ASP. For imperfective sentences, this revealed main effects of CASE for CZ ($F(1,23) = 12.04$, $p < 0.002$), CPZ ($F(1,23) = 11.56$, $p < 0.002$), PZ ($F(1,23) = 10.59$, $p < 0.003$) and POZ ($F(1,23) = 11.76$, $p < 0.002$) and for the two posterior regions (left: $F(1,23) = 8.63$, $p < 0.007$; right: $F(1,23) = 19.78$, $p < 0.001$). For perfective sentences, the main effect of CASE reached significance at FCZ ($F(1,23) = 8.88$, $p < 0.006$), CZ ($F(1,23) = 17.77$, $p < 0.001$), CPZ ($F(1,23) = 18.20$, $p < 0.001$), PZ ($F(1,23) = 12.26$, $p < 0.002$) and POZ ($F(1,23) = 8.18$, $p < 0.008$) and the two posterior regions (left: $F(1,23) = 8.63$, $p < 0.007$; right: $F(1,23) = 9.76$, $p < 0.004$).

B.2. 750–1100 ms

The statistical analysis of the 750–1100 ms time window showed main effects of CASE (midline: $F(1,23) = 31.83$, $p < 0.001$; lateral: $F(1,23) = 14.38$, $p < 0.001$) and ASP (midline: $F(1,23) = 19.43$, $p < 0.001$; lateral: $F(1,23) = 12.97$, $p < 0.001$) and interactions of CASE*ASP (midline: $F(1,23) = 18.04$, $p < 0.001$; lateral: $F(1,23) = 8.92$, $p < 0.001$) and ROI*CASE*ASP (midline: $F(5,115) = 4.23$, $p < 0.04$; lateral: $F(3,69) = 8.13$, $p < 0.001$). Resolving the interaction by ROI revealed interactions of CASE*ASP for FCZ ($F(1,23) = 10.76$, $p < 0.003$), CZ ($F(1,23) = 30.75$, $p < 0.001$), CPZ ($F(1,23) = 28.34$, $p < 0.001$), PZ ($F(1,23) = 10.98$, $p < 0.003$), POZ ($F(1,23) = 9.90$, $p < 0.004$) and for the left-posterior ($F(1,23) = 8.24$, $p < 0.008$), right-anterior ($F(1,23) = 6.04$, $p < 0.02$) and right-posterior ($F(1,23) = 16.71$, $p < 0.001$) regions.

When these interactions were further resolved by ASP, effects of CASE were observed for imperfective sentences at FCZ ($F(1,23) = 19.87$, $p < 0.001$), CZ ($F(1,23) = 50.28$, $p < 0.001$), CPZ ($F(1,23) = 49.30$, $p < 0.001$), PZ ($F(1,23) = 24.03$, $p < 0.001$), POZ ($F(1,23) = 18.49$, $p < 0.001$) and in the left-posterior ($F(1,23) = 20.75$, $p < 0.001$), right-anterior ($F(1,23) = 10.84$, $p < 0.003$) and right-posterior ($F(1,23) = 28.54$, $p < 0.001$) ROIs. For perfective aspect, there were no effects of CASE in any ROI.

Appendix C. ERP analyses relative to the onset of the aspect marker including a baseline correction

As noted in Section 3, all ERP analyses in the present paper were computed without the application of a baseline correction (for an in-depth motivation for this type of analysis, see Wolff et al., 2008). In the following, we show that the central result – namely the N400 – late positivity pattern for a violation of subject case marking – is also observable when the averages are calculated relative to a baseline (−200 to 0 ms). For the sake of brevity, we only present data relative to the onset of the aspect marker.

Fig. 6 presents grand averages at the position of the aspect marker computed relative to a −200 to 0 ms baseline. As is apparent from the figure, the ERP pattern appears highly comparable to that shown in Fig. 1, i.e. for the analysis without a baseline. This descriptive impression was confirmed by the statistical analysis, which we present in the following.

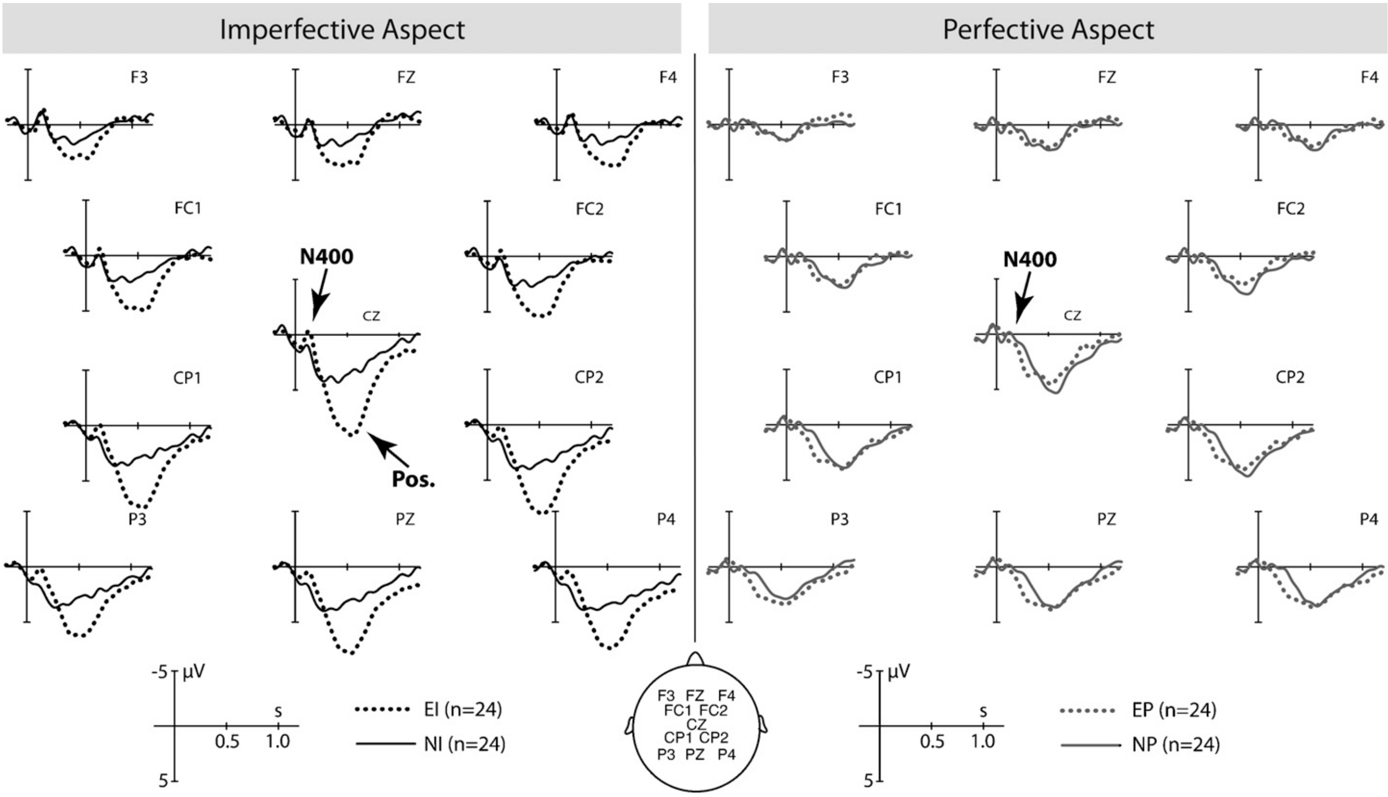

Fig. 6. Grand average ERPs ($N = 24$) time-locked to the onset of the critical aspect marker (onset at the vertical bar) and computed relative to a −200 to 0 ms baseline.

C.1. 100–300 ms

The statistical analysis showed main effects of ASP (midline: $F(1,23) = 14.10$, $p < 0.001$; lateral: $F(1,23) = 8.43$, $p < 0.008$) and interactions ROI*ASP (midline: $F(5,115) = 7.50$, $p < 0.003$; lateral: $F(3,69) = 6.72$, $p < 0.005$), CASE*ASP (midline: $F(1,23) = 5.39$, $p < 0.03$; lateral: $F(1,23) = 9.67$, $p < 0.005$) and ROI*CASE*ASP (midline: $F(5,115) = 8.03$, $p < 0.001$; lateral: $F(1,23) = 7.93$, $p < 0.001$).

Resolving the interactions by ROI showed interactions of CASE*ASP at CZ ($F(1,23) = 5.16$, $p < 0.03$), CPZ ($F(1,23) = 7.82$, $p < 0.01$), PZ ($F(1,23) = 10.61$, $p < 0.004$), POZ ($F(1,23) = 12.57$, $p < 0.002$), and at left-posterior ($F(1,23) = 18.71$, $p < 0.001$) and right-posterior ($F(1,23) = 18.33$, $p < 0.001$) sites. [N.B. effects of ASP in individual ROIs are not reported here for reasons of brevity.]

The interactions of CASE*ASP in individual ROIs were subsequently resolved by ASP. This resolution revealed simple effects of CASE for the imperfective aspect at PZ ($F(1,23) = 3.97$, $p < 0.06$), POZ ($F(1,23) = 7.16$, $p < 0.01$), and in the left-posterior ($F(1,23) = 8.36$, $p < 0.008$) and right-posterior ($F(1,23) = 7.98$, $p < 0.009$) regions. For the perfective aspect, simple effects of CASE were observed at CPZ ($F(1,23) = 4.40$, $p < 0.05$), PZ ($F(1,23) = 5.19$, $p < 0.03$), POZ ($F(1,23) = 3.82$, $p < 0.06$), and at left-posterior ($F(1,23) = 5.14$, $p < 0.03$) and right-posterior ($F(1,23) = 5.89$, $p < 0.02$) sites.

C.2. 400–700

In the second time window, the analysis showed main effects of CASE (midline: $F(1,23) = 18.76$, $p < 0.001$; lateral: $F(1,23) = 11.16$, $p < 0.003$) and ASP (midline: $F(1,23) = 20.15$, $p < 0.001$; lateral: $F(1,23) = 17.77$, $p < 0.001$) and interactions of ROI*CASE (midline: $F(5,115) = 7.64$, $p < 0.001$; lateral: $F(3,69) = 5.69$, $p < 0.006$), ROI*ASP (only at midline sites: $F(5,115) = 8.07$, $p < 0.001$), CASE*ASP ($F(1,23) = 24.88$, $p < 0.001$; lateral: $F(1,23) = 18.04$, $p < 0.001$) and ROI*CASE*ASP (midline: $F(5,115) = 5.09$, $p < 0.01$; lateral: $F(3,69) = 4.55$, $p < 0.009$).

Resolving the interactions by ROI showed an interaction of CASE*ASP at all midline sites (maximum at CZ: $F(1,23) = 32.07$, $p < 0.001$; minimum at FZ: $F(1,23) = 9.01$, $p < 0.006$) and in all lateral regions (maximum in the right-posterior ROI: $F(1,23) = 19.64$, $p < 0.001$; minimum in the left-anterior ROI: $F(1,23) = 6.40$, $p < 0.02$). [N.B. effects of CASE and ASP in individual ROIs are not reported here for reasons of brevity.]

The interactions of CASE*ASP in individual ROIs were subsequently resolved by ASP. For the imperfective aspect, this resolution revealed simple effects of CASE at all midline sites (maximum at CPZ: $F(1,23) = 55.77$, $p < 0.001$; minimum at FZ: $F(1,23) = 12.27$, $p < 0.002$) and in all lateral regions (maximum in the right-posterior region: $F(1,23) = 39.85$, $p < 0.001$; minimum in the left-anterior region: $F(1,23) = 8.90$, $p < 0.007$). There were no significant effects for the perfective aspect in any region.

References

Bartke, S., Rösler, F., Streb, J., & Wiese, R. (2005). An ERP-study of German "irregular" morphology. *Journal of Neurolinguistics*, *18*, 29–55.

Bornkessel, I., & Schlesewsky, M. (2006). The Extended Argument Dependency Model: A neurocognitive approach to sentence comprehension across languages. *Psychological Review*, *113*, 787–821.

Bornkessel, I., McElree, B., Schlesewsky, M., & Friederici, A. D. (2004). Multi-dimensional contributions to garden path strength: Dissociating phrase structure from case marking. *Journal of Memory and Language*, *51*, 495–522.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008). An alternative perspective on "semantic P600" effects in language comprehension. *Brain Research Reviews*, *59*, 55–73.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2009). The role of prominence information in the real time comprehension of transitive constructions: A cross-linguistic approach. *Language and Linguistics Compass*, *3*, 19–58.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (in press). *Processing syntax and morphology: A neurocognitive perspective*. Oxford: Oxford University Press.

Butt, M., & King, T. H. (2005). The status of case. In V. Dayal, & A. Mahajan (Eds.), *Clause structure in South Asian languages*. Berlin: Springer.
de Hoop, H., & Lamers, M. A. J. (2006). Incremental distinguishability of subject and object. In L. Kulikov, A. Malchukov, & P. de Swart (Eds.), *Case, valency and transitivity* (pp. 269–287). Amsterdam: John Benjamins.
Federmeier, K. D., & Kutas, M. (2001). Meaning and modality: Influences of context, semantic memory organization, and perceptual predictability on picture processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 202–224.
Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*(2), 78–84.
Friederici, A. D., Wang, Y., Herrmann, C. S., Maess, B., & Oertel, U. (2000). Localization of early syntactic processes in frontal and temporal cortical areas: A magnetoencephalographic study. *Human Brain Mapping*, *11*, 1–11.
Frisch, S., & Schlesewsky, M. (2001). The N400 indicates problems of thematic hierarchizing. *Neuroreport*, *12*, 3391–3394.
Frisch, S., & Schlesewsky, M. (2005). The resolution of case conflicts from a neurophysiological perspective. *Cognitive Brain Research*, *25*, 484–498.
Graddol, D. (2004). The future of language. *Science*, *303*, 1329–1331.
Hagoort, P. (2003). How the brain solves the binding problem for language: A neurocomputational model of syntactic processing. *Neuroimage*, *20*, S18–S29.
Hagoort, P. (2005). On Broca, brain, and binding: A new framework. *Trends in Cognitive Sciences*, *9*, 416–423.
Hagoort, P., & van Berkum, J. J. A. (2007). Beyond the sentence given. *Philosophical Transactions of the Royal Society B*, *362*, 801–811.
Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes*, *8*, 439–483.
Haupt, F. S., Schlesewsky, M., Roehm, D., Friederici, A. D., & Bornkessel-Schlesewsky, I. (2008). The status of subject–object reanalyses in the language comprehension architecture. *Journal of Memory and Language*, *59*, 54–96.
Hoeks, J. C. J., Stowe, L. A., & Doedens, G. (2004). Seeing words in context: The interaction of lexical and sentence level information during reading. *Cognitive Brain Research*, *19*, 59–73.
Huynh, H., & Feldt, L. S. (1970). Conditions under which the mean-square ratios in repeated measurement designs have exact *F*-distributions. *Journal of the American Statistical Association*, *65*, 1582–1589.
Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language*, *52*, 205–225.
King, J. W., & Kutas, M. (1995). Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience*, *5*, 196–214.
Kluender, R., & Kutas, M. (1993). Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience*, *5*, 196–214.
Kolk, H. H. J., Chwilla, D. J., van Herten, M., & Oor, P. J. (2003). Structure and limited capacity in verbal working memory: A study with event-related potentials. *Brain and Language*, *85*, 1–36.
Kuperberg, G. R., Sitnikova, T., Caplan, D., & Holcomb, P. (2003). Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research*, *17*, 117–129.
Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*, 463–469.
Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*, 203–205.
Kutas, M., Van Petten, C., & Kluender, R. (2006). Psycholinguistics electrified II (1994–2005). In M. Traxler, & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 659–724). London: Elsevier.
Lamers, M. A. J., & de Hoop, H. (2005). Animacy information in human sentence processing: An incremental optimization of interpretation approach. *Lecture Notes in Artificial Intelligence*, *3438*, 158–171.
Lee, H. (2003). Prominence mismatch and markedness reduction in word order. *Natural Language and Linguistic Theory*, *21*, 617–680.
McClelland, J. L., & Patterson, K. (2002). Rules or connections in past-tense inflections: What does the evidence rule out? *Trends in Cognitive Sciences*, *6*, 465–472.
Mohanan, T. (1994). *Argument structure in Hindi*. Stanford, CA: CSLI Publications.
Nevins, A., Dillon, B., Malhotra, S., & Phillips, C. (2007). The role of feature-number and feature-type in processing Hindi verb agreement violations. *Brain Research*, *1164*, 81–94.
Nigam, A., Hoffman, J., & Simons, R. (1992). N400 to semantically anomalous pictures and words. *Journal of Cognitive Neuroscience*, *4*, 15–22.
Osterhout, L., & Holcomb, P. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*, 785–806.
Philipp, M., Bornkessel-Schlesewsky, I., Bisang, W., & Schlesewsky, M. (2008). The role of animacy in the real time comprehension of Mandarin Chinese: Evidence from auditory event-related brain potentials. *Brain and Language*, *105*, 112–133.
Pinker, S., & Ullman, M. T. (2002). The past and future of the past tense. *Trends in Cognitive Sciences*, *6*, 456–463.
Rietveld, A. C. M., & Gussenhoven, C. (1985). On the relation between pitch excursion size and prominence. *Journal of Phonetics*, *13*, 299–308.
t'Hart, J., Collier, R., & Cohen, A. (1990). *A perceptual study of intonation*. Cambridge: Cambridge University Press.
Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, *92*, 231–270.
van den Brink, D., & Hagoort, P. (2004). The influence of semantic and syntactic context constraints on lexical selection and integration in spoken-word comprehension as revealed by ERPs. *Journal of Cognitive Neuroscience*, *16*, 1068–1084.
van Herten, M., Chwilla, D. J., & Kolk, H. H. J. (2006). When heuristics clash with parsing routines: ERP evidence for conflict monitoring in sentence perception. *Journal of Cognitive Neuroscience*, *18*, 1181–1197.
Vissers, C. T. W. M., Kolk, H. H. J., van de Meerendonk, N., & Chwilla, D. J. (2008). Monitoring in language perception: Evidence from ERPs in a picture-sentence matching task. *Neuropsychologia*, *46*, 967–982.
Weyerts, H., Penke, M., Dohrm, U., Clahsen, H., & Münte, T. F. (1997). Brain potentials indicate differences between regular and irregular German plurals. *Neuroreport*, *8*, 957–962.
Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience*, *16*, 1271–1288.
Wolff, S., Schlesewsky, M., Hirotani, M., & Bornkessel-Schlesewsky, I. (2008). The neural mechanisms of word order processing revisited: Electrophysiological evidence from Japanese. *Brain and Language*, *107*, 133–157.
Ye, Z., Zhan, W., & Zhou, X. (2007). The semantic processing of syntactic structure in sentence comprehension: An ERP study. *Brain Research*, *1142*, 135–145.