An ERP Study of Feature-Matching in Hindi Verb Agreement

Andrew Nevins[1]
Brian Dillon[2]
Colin Phillips[2,3]

[1]*Department of Linguistics, Harvard University,* [2]*Department of Linguistics, University of Maryland, College Park* [3]*Neuroscience and Cognitive Science Program, University of Maryland, College Park*

Total number of pages: 28
Number of tables: 6 (incl. appendix)
Number of figures: 2

Address for correspondence:

Andrew Nevins
Department of Linguistics
Harvard University
317 Boylston Hall
Cambridge, MA 02138

nevins@fas.harvard.edu
(617) 495-8107 (ph.)
(617) 496-4447 (fax)

Preparation of this paper was supported by grants to CP from the National Science Foundation (BCS-0196004) and the Human Frontier Science Program (RGY-0134) and by a Semester Research Award from the University of Maryland. We are very grateful to Shradha Upadhyay for assistance in constructing materials and recruiting participants and to Alison Austin for assistance with data collection and analysis.

Abstract

This article presents an event-related potentials (ERP) study of Hindi that investigates whether electrophysiological responses to syntactic agreement violations vary as a function of the type and number of incorrect agreement features. Hindi is well suited to investigation of this issue, since verbs in Hindi mark agreement with the person, number, and gender features of the subject noun phrase. Evoked responses were recorded for visually presented verbs appearing at the end of a sentence-initial adverbial clause, comparing responses in a grammatically correct condition with four grammatically incorrect conditions that mismatched the correct agreement on different dimensions (gender, number, gender+number, person+gender). A P600 response was elicited in all grammatically incorrect conditions. No amplitude differences were found among the gender, number, and gender+number conditions. This suggests that the degree of overlap between observed and expected word forms at the morphosyntactic level does not impact ERP responses, contrasting with findings at the semantic level (Federmeier & Kutas, 1999) and in auditory processing (Paavilainen et al., 2001). A significantly larger P600 response was elicited in the person+gender condition, an effect that is attributed to the greater salience of the person feature in multiple levels of representation.

Theme: Neural basis of behavior

Topic: Cognition

Keywords: Event-related potentials, sentence processing, agreement, P600, Hindi

1. Introduction

A leading result from research on language using event-related brain potentials (ERPs) has been the documentation of a series of temporally and topographically distinct response components that are elicited by different kinds of unexpected linguistic material. Words that are syntactically appropriate but semantically inappropriate characteristically elicit a central negativity known as the N400 (Kutas & Hillyard, 1980; Kutas & Federmeier, 2000). Infrequent oddball syllables that interrupt an otherwise uniform train of sounds elicit a response component with a 150-300 ms latency known as the mismatch negativity (MMN: Näätänen, Gaillard, & Mäntysalo, 1978; Näätänen, Tervaniemi, Sussman, Paavilainen, & Winkler, 2001). Words that are morphologically or syntactically incorrect elicit a late positivity known as the P600 (Friederici, Pfeifer, & Hahne, 1993; Hagoort, Brown, & Groothusen, 1993; Osterhout & Holcomb, 1992), and in some cases also an (early) left anterior negativity ((E)LAN: Coulson, King, & Kutas, 1998; Friederici et al., 1993; Hagoort, Wassenaar, & Brown, 2003; Neville, Nicol, Barss, Forster, & Garrett, 1991). Since the initial discoveries of these response components there has been substantial interest in the goal of understanding what specific factors control the timing and amplitude of these effects. This interest has been guided by the assumption that an improved understanding of this variation will lead to more precise models of the representations and processes that underlie the ERP effects.

The aim of the current study is to use the rich verbal agreement system of Hindi to investigate the mechanisms that are used to anticipate and/or check syntactic agreement relations. Future tense verbs in Hindi are marked for person, gender and number agreement with the subject of the clause. This contrasts with languages like Spanish, German, and English, in which verbal agreement rarely encodes gender information and uses a single agreement morpheme to encode both person and number information. The current study investigates ERP responses to various correct and incorrect subject-verb agreement configurations in Hindi in order to assess whether ERPs elicited by agreement

violations vary as a function of the nature and/or the number of incorrect agreement features.

Quantitative variation in ERP responses elicited by unexpected linguistic material has been investigated in detail in the domains of semantic interpretation and auditory processing, and these therefore serve as a useful point of comparison to the morphosyntactic phenomena investigated here. Previous research on semantics and auditory perception has shown that qualitative variation in ERP response amplitudes is associated with at least two dimensions of the relation between expected and unexpected material.

Mismatches between expected and unexpected/incongruous linguistic items may be quantified in terms of relative probabilities of occurrence, or in terms of the degree of overlap (or 'matching') in properties between the expected and unexpected items. In studies of semantic processing much evidence indicates that N400 amplitude varies as a function of the probability that the target word should appear given its preceding context. For example, following the context *She was stung by a ...* words that are implausible or anomalous, such as *fly* or *dog*, are clearly less probable than the word *bee*, and thus elicit larger N400 responses (Kutas & Hillyard, 1980). Furthermore, a final word that is plausible but less probable, such as *wasp*, also elicits a larger N400 response than the most probable completion. Thus, N400 amplitude is considered to inversely correlate with the *cloze probability* of a syntactically appropriate content word (Kutas & Hillyard, 1984; van Petten & Kutas, 1990, 1991). In the domain of syntax there is some evidence for similar variation in P600 amplitude as a function of the probability of different types of arguments following a verb (Osterhout, Holcomb, & Swinney, 1994).

Recent evidence suggests that N400 amplitude may also vary as a function of the degree of semantic overlap, as quantified by a feature-space, between a target word and the expected word, independent of cloze probability. A study by Federmeier & Kutas (1999) varied the semantic properties of sentence-final words in highly constraining contexts in an attempt to investigate the role of memory organization in linguistic processing. They found that pairs of words that were entirely implausible and improbable

but varied in their semantic feature overlap with the expected word elicited different amplitude N400 responses, with a reduced N400 in cases of greater semantic overlap. In (1), for example, the word *pines* elicited a smaller N400 than *tulips*, relative to the expected final word, *palms*. Federmeier & Kutas suggest that this feature-overlap advantage reflects prediction of the expected word based on context, which in turn leads to priming of the semantic features of words that are inappropriate yet related to the expected word.

(1) They wanted to make the hotel look more like a tropical resort. So, along the driveway they planted rows of {palms, pines, tulips}.

Evidence for a similar feature-overlap advantage has been found in ERP studies of auditory processing. For example, Levänen and colleagues showed in an auditory mismatch paradigm that when the deviant sound mismatched the standard sound along two dimensions (frequency and duration) a magnetic mismatch field (MMF) was elicited that was roughly the sum of the individual MMFs elicited by either mismatched dimension individually (Levänen, Hari, McEvoy, & Sams, 1993). These findings were extended in a subsequent study in which stimuli included two- and three-feature deviants, manipulating the dimensions of frequency, intensity, and stimulus-onset asynchrony (Paavilainen, Valppu, & Näätänen, 2001). This study found that MMN amplitudes increased with greater featural distance between the standard and deviant sounds, although amplitudes were additive for 2-feature mismatches and less than additive for 3-feature mismatches. These findings are compatible with a standard account of the auditory mismatch response as a reflection of a process of comparison between a deviant stimulus and a memory trace of the standard sound (Näätänen, 1992).

Thus we find evidence in at least two domains for feature priming, understood as a process in which the features of a likely upcoming item are primed. By hypothesis, priming can be detected based upon reduced amplitude of the MMN or N400, which implies reduced neural activity, and hence reduced processing requirements. The goal of this study is to investigate whether a similar feature-overlap advantage is found in the

domain of morphosyntax. By manipulating the degree of overlap between correct and incorrect agreeing verb forms it is possible to ask whether agreement violations are processed in a manner that reflects degrees of deviance, or whether agreement violation detection is an 'all or nothing' phenomenon. If the processing of agreement violations parallels findings in semantic processing and auditory perception described above, then we should expect that a verb agreement violation that shares at least some features with the target form should be easier to process than a more deviant form. A second question of interest is whether all agreement features are analyzed in the same manner, or whether an agreement mismatch in one feature (e.g. person) is analyzed as a more severe deviation than a mismatch in another feature (e.g. number).

There are three prerequisites for a study of this kind: one or more electrophysiological components that are sensitive to agreement violations, independently motivated evidence for a notion of 'degree of overlap' in morphosyntax, and a language that has appropriate morphology for conducting this test.

Many previous studies have documented the characteristic ERP responses associated with agreement violations. A common finding is that verb inflection errors involving agreement or tense elicit a biphasic response consisting of an anterior negativity that is sometimes left-lateralized (LAN) followed by P600 response with a broad posterior scalp distribution (Coulson et al., 1998; Friederici et al., 1993; Gunter, Stowe, & Mulder, 1997; Hagoort & Brown, 2000; Kaan 2002; Kutas & Hillyard, 1983; Münte, Matzke, & Johannes, 1997; Osterhout & Mobley, 1995), although some studies report that verb inflection errors elicit only the P600 response (Gunter & Friederici, 1999; Hagoort et al 1993; Lau, Stroud, Plesch, & Phillips, in press; Osterhout & Nicol, 1999). Studies of subject-verb agreement processing have typically focused on person and number features, since these are more commonly encoded on the verbs of well-studied European languages. Studies of gender agreement processing have involved relations between nouns and determiners or adjectives. Gender agreement violations have been found to elicit familiar P600 effects (Hagoort & Brown, 1999) or biphasic LAN-P600 patterns (Gunter, Friederici, & Schriefers, 2000; Barber & Carreiras, 2005). In sum, current

evidence suggests that agreement violations involving person, number, and gender features elicit broadly similar ERP responses, with variation in the responses elicited by different features restricted to quantitative differences in the timing and amplitude of the P600 (Barber & Carreiras, 2005).

There is some independent motivation for a notion of degree of overlap or 'matching' in agreement features. In discussions of paradigm structure in morphological theory it has been argued that the closeness of cells in a morphological paradigm is an important factor in determining syncretism, i.e., the possibility of a single morpheme being used for multiple cells in the paradigm (e.g., Williams, 1994; Wunderlich, 1996). However, such claims leave open the question of whether the distance between a pair of distinct morphemes in a morphological paradigm maps onto graded measures of incongruence or surprise in a way that predicts differences in ERP responses.

It should be noted that it is only useful to quantify the degree of overlap or 'feature distance' between a pair of words if the features of the words are analyzed as a group. In the domain of word meaning this is a largely uncontroversial assumption, and hence the predictions for N400 amplitude variation are relatively straightforward. In contrast, it is less clear whether the features of richly inflected verbs (e.g., person, number, gender, tense) are analyzed jointly or individually. Much research in syntactic theory (e.g., Chomsky, 1995; Jackendoff, 2002) assumes implicitly that inflectional features are analyzed jointly. If, however, mismatches in distinct agreement morphemes are analyzed individually and elicit independent ERP responses, then we should expect these responses to sum, leading to larger ERP responses to forms that mismatch in multiple agreement features. Consequently, care must be exercised in drawing inferences about processing mechanisms from ERP responses elicited by agreement mismatches, since summed responses are compatible either with individual analysis of each feature, or with a feature distance metric operating on feature-bundles. Relatedly, analyses of the detailed responses to morphosyntactic violations do not bear on recent debates about the specificity of ERP response components such as the P600 to linguistic or syntactic

processes (e.g., Coulson et al., 1998; Osterhout & Hagoort, 1999; Patel, Gibson, Ratner, Besson, & Holcomb, 1998).

Hindi future tense morphology is well suited to exploration of these questions, since it preserves a complete paradigm in which all future tense verb forms are inflected for person, number, and gender on a single word. This is rare among well-studied languages, which typically separate person and number onto an auxiliary, and number and gender onto a participle (indeed, Hindi does this in non-future tenses). This property of Hindi makes it possible to place all three independent feature-value mismatches within the same word. We varied the number and nature of feature mismatches across experimental conditions, which included two 1-feature violations (gender, number), and two 2-feature violations (gender/number, person/gender). (Note that a person/number violation could not be presented in this paradigm due to a syncretism between 1st person plural and 3rd person plural agreement.)

Hindi future tense morphology follows a regular inflectional paradigm. Some example future tense verb forms are shown in Table 1, which contains the forms in *Devanagari* script, its phonetic form, a morpheme-by-morpheme gloss, and a translation. Future tense inflection consists of two syllables. Person is marked exclusively on the first syllable, number is marked on both syllables, and gender is marked on the last syllable. A more detailed explanation of the relevant morphological forms, as well as their phonetic and orthographic representations, can be found in Appendix A.

गायेगा			गायेंगे		
	gaayegaa			**gaayengee**	
gaay-	**e-**	**gaa**	**gaay-**	**en-**	**gee**
sing	*3rd/Singular*	*Masc/Singular*	*sing*	*3rd/Plural*	*Masc/Plural*
"He will sing"			"They (masc.) will sing"		

Table 1: Hindi verb agreement forms, future tense

In summary, the current study makes use of the flexible Hindi verb agreement paradigm in order to investigate whether morphosyntactic processing shows the same

feature-overlap advantage observed in other domains, including auditory processing and higher-level semantic processing. ERP is an appropriate technique, since it does not require any task beyond reading for comprehension, and allows direct inferences about the timing of processes with millisecond-level resolution. These results in turn will shed light on questions specific to morphosyntax about the relative salience of inflectional features.

2. Methods

2.1 Participants

Twenty-three members of the University of Maryland community participated in this study. Data from four participants were excluded due to unacceptably high levels of artifacts in the EEG recordings. The remaining 19 participants (6 females; mean age 23.9 years) were all healthy, native speakers of standard Hindi with no history of neurological disorder, and all were strongly right-handed based on the Edinburgh Handedness Inventory (Oldfield, 1971). In order to screen for mastery of standard Hindi agreement morphology and fluency in reading the Hindi *Devanagari* script, all participants took part in an off-line pre-test, consisting of 15 questions that addressed possible variation in grammatical forms. A number of speakers of non-standard Hindi dialects were excluded based on errors in this pre-test, and a small number of additional participants were excluded because they lacked the reading fluency needed to comprehend Hindi sentences presented in an RSVP paradigm. All participants gave informed consent and were paid $10/hour for their participation, which lasted around 2.75 hours, including set-up time.

2.2 Design and Materials

The experiment manipulated the congruency of subject verb agreement across 5 conditions, taking advantage of the fact that Hindi future tense verbs are marked separately for person, gender, and number. The critical verb was the final word of a

sentence-initial adverbial clause, thereby reducing the risk of ERP effects associated with sentence-final wrap-up. A grammatical control condition with 3rd person masculine singular verb agreement was contrasted with four ungrammatical conditions, consisting of two conditions that mismatched in just one agreement feature (gender, number) and two conditions that mismatched in two agreement features (gender/number, person/gender). We tested only 4 of the 7 possible combinations of ungrammatical agreement features due to constraints on the length of the experiment, and because combined person/number violations are not possible, due to syncretism (i.e., identical forms) between 1st person plural and 3rd person plural. All conditions were identical except for the agreement suffixes on the critical verb.

The critical verb always appeared as the sixth word of the sentence and either matched or mismatched with a nominative subject noun phrase (NP) that appeared in word positions 2-4. The first word of the subject NP was a demonstrative determiner, which is both natural and common in Hindi, due to the lack of definite determiners. The demonstrative was followed by an adjective-noun sequence that was distinctively marked as masculine singular on at least one word. Nominative masculine nouns and adjectives in Hindi follow one of two patterns, either with number specific variants, e.g., *rasoiyaa / rasoiyee,* 'cook' sg/pl., *dublaa/dublee,* 'thin' sg./pl., or with number invariant forms, e.g., *jaj,* 'judge', sg. or pl., *samajhdaar*, 'sensible' sg. or pl. In all experimental items at least one of the adjective or the noun was distinctively marked as singular. Hindi has canonical verb-final word order, and consequently the critical verb was placed inside a sentence-initial adverbial clause, such that it would not appear in sentence-final position.

150 sets of 5 conditions were distributed across 5 lists in a Latin Square design. Three sentence-initial subordinators each appeared in 50 sets of items, *haalanki*, 'although', *chunki*, 'since', i.e., 'due to the fact that...', *jab*, 'when', i.e., 'at the time that...'. Thus, participants read 30 sentences with correct agreement and 120 sentences with incorrect agreement. The filler items included 120 sentences in a past-tense frame, 60 of which displayed incorrect tense on the embedded verb, and 180 additional filler items using the same subordinators as above. 45 of the filler items displayed various noun-phrase internal

agreement errors. Thus, each sentence list contained a total of 225 correct and 225 incorrect sentences. A sample item is shown in Table 2, along with the five verb forms corresponding to each of the five conditions, in both Roman and *Devanagari* script. Each condition is differentiated from the others by vowel diacritics found above the symbols' line, except for the person+gender violation, which is the only of the five conditions which marks the vowel change by changing the penultimate character in addition to vowel diacritics. A more detailed description of how agreement morphology is conveyed in *Devanagari* script can be found in Appendix A.

Haalanki vo safiraa sangitkaar gaanaa **gaayegaa** lekin Shrotaa hoslaa nahii baRhaayenge
although that crazy.masc musician song ***sing.fut.3rd.sng.masc*** *but morale listeners NEG enhance*
"Although that crazy musician will sing a song, the listeners' morale won't increase."

गायेगा	गायेगी	गायेंगे	गायेंगी	गाऊंगी
gaayegaa Correct Agreement	**gaayegii** Gender Violation	**gaayengee** Number Violation	**gaayengii** Gender + Number Violation	**gaayuungii** Person + Number Violation

Table 2: Sample set of experimental items for the ERP study. A Latin square design employed all 5 variants of the sentence above, differing only in the inflected verb shown in boldface for each condition.

2.3 Procedure

Participants were comfortably seated in a dimly lit testing room around 100 cm in front of a computer monitor. Sentences were presented one word at a time in black letters on a white screen in the 30 pt Devanagari font Shusha. Each sentence was preceded by a fixation cross. Participants pressed a button to initiate presentation of the sentence, which began 1000 ms later. Each word appeared on the screen for 400 ms, followed by 200 ms of blank screen. Pre-testing showed that the 600 ms SOA was a comfortable reading rate for native speakers of Hindi, and that faster presentation rates led to substantial difficulty in many participants. The last word of each sentence was marked with a period, and 1000 ms later a question mark prompt appeared on the screen. Participants were instructed to

read the sentences carefully without blinking and to indicate with a button press whether the sentence was an acceptable Hindi sentence. Feedback was provided for incorrect responses. Each experimental session was preceded by a 12 trial practice session that included both grammatical and ungrammatical sentences. Participants received feedback and were able to ask clarification questions about the task at this time. The experimental session itself was divided into six blocks of 75 sentences each. Breaks were permitted after each block, as necessary.

2.4 EEG recording

EEG was recorded from 30 Ag/AgCl electrodes, mounted in an electrode cap (Electrocap International): midline: Fz, FCz, Cz, CPz, Pz, Oz; lateral: FP1/2, F3/4, F7/8, FC3/4, FT7/8, C3/4, T7/8, CP3/4, TP7/8, P4/5, P7/8, O1/2. Recordings were referenced online to the linked average of the left and right mastoids. Additional electrodes were placed on the left outer canthus, and above and below the left eye to monitor eye movements. EEG and EOG recordings were amplified and sampled at 1 kHz using an analog bandpass filter of 0.1-70 Hz. Impedances were kept below 5 kΩ.

2.5 EEG analysis

All analyses are based upon grand averages of 1100 ms intervals surrounding the critical verb, consisting of a 100 ms baseline interval and a 1000 ms post-stimulus interval. Trials with ocular and other large artifacts were rejected based on visual screening. Among the 19 included participants, the total rejection rate was 15.8% ranging from 14.6%-17.8% across conditions. A 10 Hz low-pass filter was applied to the grand average ERPs for display purposes, however all statistics were performed on unfiltered data. ANOVAs were calculated based on mean voltages within a series of 200 ms time intervals that allowed continuous tracking of the evolution of any ERP responses elicited by the target word (0-200 ms, 200-400 ms, 400-600 ms, 600-800 ms, 800-1000 ms), plus

one additional time interval that was included based on reports of a left anterior negativity (LAN) in previous literature (300-500 ms).

For statistical analyses, six regions of interest (ROIs) were defined, consisting of three electrodes at each ROI: left anterior (FT7, F3, FC3), midline anterior (FZ, FCZ, CZ), right anterior (F4, FC4, FT8), left posterior (TP7, CP3, P3), midline posterior (CPZ, PZ, OZ), and right posterior (CP4, P4, TP8). ANOVAs were performed hierarchically, using the within-subjects factors *condition*, *anterior-posterior*, and *laterality* (left/midline/right), with follow-up analyses based on comparisons between pairs of conditions. All *p*-values reported below reflect the application of the Greenhouse-Geisser correction where appropriate to control for violations of the sphericity assumption (Greenhouse & Geisser, 1959), together with the original degrees of freedom. Due to the large number of possible interactions in this design, we report as significant only those interactions for which follow-up analyses yielded significant contrasts within the levels of the interacting factors.

3. Results

3.1 Acceptability Judgment Task

Overall accuracy on the acceptability judgment task was 92%, with individual condition means as follows: grammatical control 90%, incorrect gender 91%, incorrect number 84%, incorrect gender+number 97%, incorrect person+gender 98%. The low accuracy in the number condition was primarily due to two participants, and the mean accuracy of the remaining participants on the number condition was 91%.

3.2 ERPs

Figure 1 shows topographic scalp maps of the mean difference between each violation condition and the control condition at successive 200 ms intervals. The grand averaged waveforms for all five conditions at the midline electrode CZ is shown in

Figure 2. Visual inspection indicates that all four agreement violation conditions elicited a broadly distributed posterior positivity with a peak amplitude in the 600-800ms interval, and that this positivity had an earlier onset and larger amplitude in the person+gender violation condition than in the three other violation conditions. The P600 effect continues to the 800-1000 ms interval, albeit with a reduced amplitude. Visual inspection provided no evidence of the left anterior negativity (LAN) response that has been observed in many ERP studies of syntactic violations, and this was confirmed by statistical analyses at the 300-500ms interval that is characteristic of this response. Therefore, in what follows we report ANOVA results only for the 200 ms intervals shown in Figure 1.

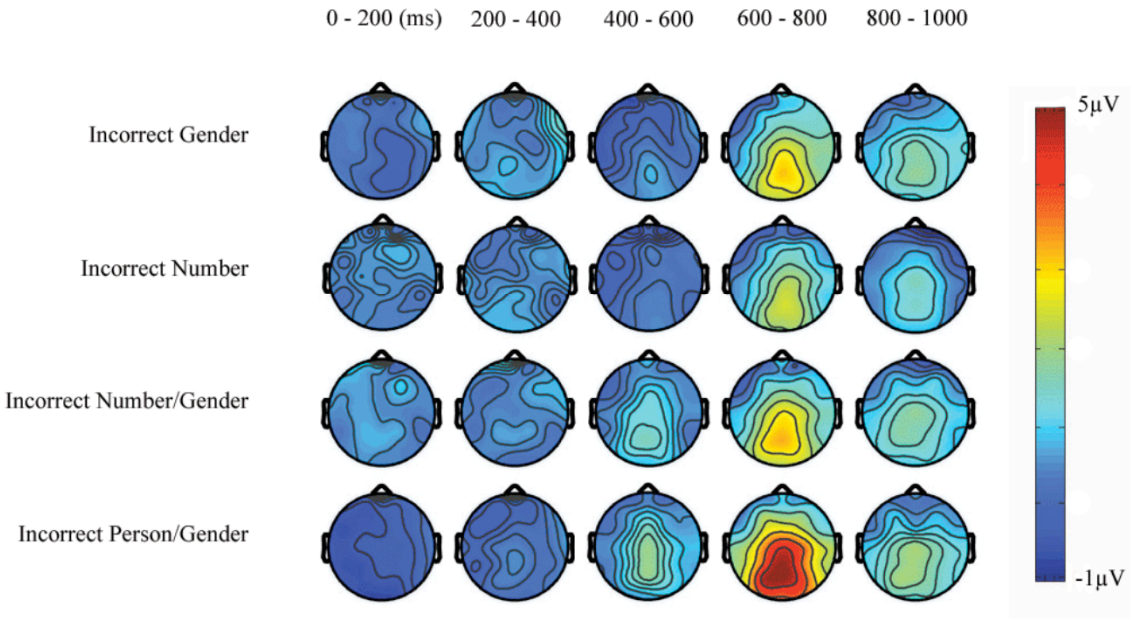

Figure 1: Topographic scalp voltage maps, showing the grand average difference between each grammatically incorrect condition and the control condition at each successive interval following the critical verb.

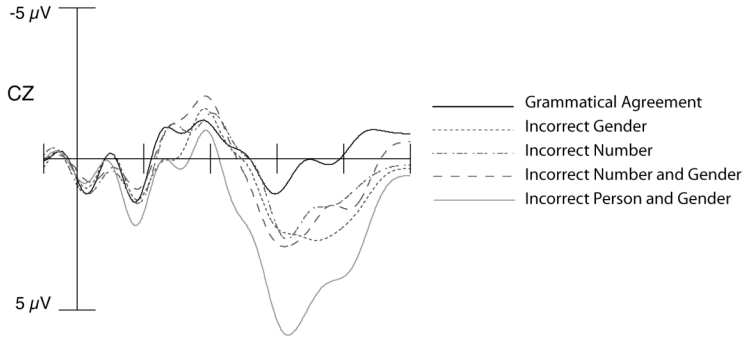

Figure 2: ERP waveforms showing grand average responses elicited by the critical verb in all 5 conditions at central electrode CZ.

In the statistical comparison of the responses to the critical verb across all five conditions no reliable effects involving the condition factor were observed at the 0-200 ms interval or the 200-400ms interval.

At the 400-600 ms interval the overall ANOVA showed no main effect of condition, but showed a three-way interaction between the factors condition, anteriority, and laterality, $F(8,144) = 2.83$, $p < .05$, reflecting the fact that the P600 response to agreement violations started in this interval and had a central posterior focus. Subsequent pairwise comparisons among conditions showed a significant main effect of condition in the comparison of the person+gender violation with each of the four other conditions (person+gender vs. control, $F(1,18) = 6.35$, $p < .05$; person+gender vs. number+gender, $F(1,18) = 6.15$, $p < .05$; person+gender vs. number, $F(1,18) = 6.24$, $p < .05$; person+gender vs. gender, $F(1,18) = 3.39$, $p < .1$. There were no other reliable differences between conditions at this interval.

At the 600-800 ms interval the overall ANOVA showed a main effect of condition, $F(4,72) = 11.01$, $p < .0001$, as well as interactions of condition and anteriority, $F(4,72) = 31.95$, $p < .0001$, and of condition and laterality, $F(8,144) = 3.66$, $p < 001$. These interactions reflect the posterior central focus of the P600 effect. Table 3 shows the results of pairwise comparisons among each condition, using the same three factors. The table shows that the main effect of condition and the condition x anteriority interaction and the condition x laterality interaction were significant for the comparison of each

violation condition with the control condition, and also for the comparison of the person+gender condition with all other violation conditions. Table 4 shows results from comparisons of the same conditions, separated into anterior and posterior regions in a two-way ANOVA with the factors condition and laterality. This analysis confirms that all mismatching conditions showed similar topographic distributions, with a posterior focus for the P600, as reflected in Table 4 in generally larger *F*-values for the effect of condition at posterior regions.

Incorrect Gender	8.451** 12.528** 4.590* -			*cond* (1,18) *cond x ant* (1,18) *cond x lat* (2,36) *cond x ant x lat* (2,36)
Incorrect Number	5.262* 7.298* 3.945* -	- - - -		
Incorrect Number/Gender	10.231** 12.044** 3.124† -	- - - -	- - - -	
Incorrect Person/Gender	39.319*** 65.209*** 9.869*** -	12.678** 4.405* 3.609* -	15.054** 9.767** 5.103* -	26.521*** 17.950*** 5.604* -
	Control	Incorrect Gender	Incorrect Number	Incorrect Number/Gender

Table 3: Results of pairwise comparisons of ERPs elicited by all conditions at 600-800 ms after the critical verb, all regions combined.

Control		3.805† 3.244†	- 3.706*	4.103† 3.505*	22.995*** 9.264**
Incorrect Gender	13.415** 5.130*		- -	- -	7.413* 2.932†
Incorrect Number	7.754* 2.871†	- -	*cond* (1,18) *cond x lat* (2,36)	- -	8.099* 5.140*
Incorrect Number/Gender	16.734** 3.359*	- -	- -		17.992*** 4.752*
Incorrect Person/Gender	55.963*** 6.736**	17.619* 3.773*	22.948*** 3.670*	33.552*** 4.737*	
Anterior / Posterior	Control	Incorrect Gender	Incorrect Number	Incorrect Number/Gender	Incorrect Person/Gender

Table 4: Results of pairwise comparisons of ERPs elicited by all conditions at 600-800 ms after the critical verb, separating anterior and posterior regions.

At the 800-1000ms interval the overall ANOVA showed a significant main effect of condition, $F(4,72) = 2.95$, $p < .05$, and a marginally significant interaction of condition and anteriority, $F(8,144) = 2.39$, $p < .1$. Pairwise comparisons revealed a significant main effect of condition in the comparison of person+gender vs. control, $F(1,18) = 7.95$, $p < .05$, person+gender vs. number+gender, $F(1,18) = 5.1$, $p < .05$, person+gender vs. number, $F(1,18) = 4.35$, $p < .1$, and gender vs. control, $F(1,18) = 5.05$, $p < .05$. As in the previous intervals, the effects were strongest in the posterior midline region.

4. Discussion

Previous ERP research on sentence comprehension has shown that the N400 response is sensitive to the degree of semantic overlap between a presented word and an expected word (Federmeier & Kutas, 1999), and similar sensitivity to feature overlap has been observed in studies of acoustic processing using the mismatch negativity response (Paavilainen et al., 2001). In both domains a feature-overlap advantage has been

observed, such that target items that are more similar to an expected item elicit smaller ERP responses. The current experiment investigated whether a similar effect is found in the domain of morphosyntax, taking advantage of the rich verbal agreement paradigm of Hindi to test four different types of agreement violation. All four agreement violations elicited a positivity with the characteristic latency and posterior scalp distribution of the P600 response. The most direct test of the feature-matching effect for agreement showed no evidence of smaller P600 responses for violations that were more similar to the expected form: the response to a combined number+gender violation showed the same amplitude as the response to number and gender violations individually. We discuss below several possible reasons for the lack of a feature overlap advantage.

A greater amplitude P600 response was observed in the person+gender violation condition than in the other agreement violation conditions. However, it is likely that this effect was due to a strong response to violation of the person feature, rather than to the additive combination of responses to two violations. Below we discuss converging evidence for the special status of person in grammatical agreement systems.

The statistical analysis indicated that the P600 effect persisted until the 800-1000 ms interval in the gender violation condition but not in the number violation condition. This difference in the duration of the response is consistent with results from a related gender vs. number comparison in Spanish (Barber & Carreiras, 2005). However, this difference was numerically small, as can be seen from Figure 2, and it does not impact the overall conclusions of this study, and therefore we do not discuss this difference further.

4.1 Lack of a Feature Distance Effect

The lack of a difference between the P600 responses to violations of number, of gender, and of number and gender combined is interesting in light of other evidence for systematic variation in the amplitude of ERP responses elicited by semantic and syntactic anomalies. In addition to the many studies that have reported systematic amplitude variation in the N400 response (Kutas & Hillyard, 1984, van Petten & Kutas, 1991,

Titone & Salisbury, 2004), some studies have reported systematic amplitude variation in P600 responses. Osterhout and colleagues found that the P600 elicited by garden path sentences changes as a function of verb subcategorization biases (Osterhout, Holcomb, & Swinney, 1994). Other studies have reported P600 modulation to be a function of the saliency of a morphosyntactic violation (Coulson, King & Kutas 1998), as well as covarying with proficiency level in second language learners (Hahne, 2001). In the current study we also found P600 amplitude variation, distinguishing the person+gender violation from the other violations. We can therefore exclude the possibility that the lack of a feature overlap advantage for the P600 is simply due to the fact that the P600 fails to show amplitude variation in general.

A number different possible accounts of the lack of a feature overlap advantage present themselves. A first possibility is that there is a qualitative difference between the representation of morphosyntactic features and the representation of semantic or acoustic features, with the specific consequence that the degree of overlap between an expected word and a target verb form does not affect P600 amplitudes. This could be because morphosyntactic features are formal linguistic features, whereas semantic and acoustic features are grounded in language-independent cognitive representations. Alternatively, this could reflect the fact that verb agreement is an all-or-nothing phenomenon. If a verb does not match the agreement features of its subject, then this grammatical relation is simply not reflected, and thus there is no benefit to an approximately correct form of agreement. Relatedly, there are differences in accuracy in speech production between semantics, acoustics, and morphosyntax: substitution of open-class words with semantic associates is not uncommon, and speech production tolerates a good amount of acoustic variability. In contrast, errors in subject-verb agreement are rare in speech production in unimpaired adults.

A second possible reason is that the priming mechanism that is plausibly responsible for the feature overlap advantage in semantic and acoustic processing is not implicated in morphosyntactic processing. Federmeier and Kutas (1999, p. 387) suggest that the semantic features of a highly expected word are ‘pre-activated’ based on contextual cues,

and that this has the consequence that words that share features with the pre-activated word are partially primed, and hence elicit smaller N400 responses. Therefore, the lack of a feature overlap advantage for morphosyntax might reflect the absence of prediction or a lack of priming of morphosyntactic features. An absence of agreement priming could, in principle, reflect an absence of prediction in morphosyntactic processing in general. An alternative that we consider more likely is that morphosyntactic features might not have been strongly predicted at the verb position in our study because the specific position of the verb could not be reliably anticipated. As shown in Table 2 the critical verb always appeared following a direct object NP. Although it is natural for a verb to follow a direct object NP in Hindi, other categories such as adverbs and benefactive phrases may appear following a direct object NP, and therefore the parser might not have consistently anticipated the appearance of a verb following the direct object NP. Perhaps even more important is the fact that the parser might not have consistently predicted the actual phonological form of agreement features that will be marked on a verb form, and whether or not they will be synthetically realized within a single form, or analytically realized with a participle and an auxiliary. For example, the future tense in Hindi is marked by a single synthetic form which realizes person, gender and number, but the present tense is marked by a participle that realizes number and gender and an auxiliary that realizes person and number, while the simple past tense is marked by a participle that marks gender and number with no auxiliary and no expression of person. Prior to presentation of the verb, however, our experimental conditions were compatible with a variety tenses, and thus participants would not be able to predict the specific form of the agreement morphemes. The lack of a feature overlap advantage could thus implicate a lack of priming of morphosyntactic features, although this might reflect a specific property of Hindi rather than a property of morphosyntax in general.

4.2 Why Person Is Stronger

Although our results indicate that the P600 elicited by a number+gender violation is no larger than the P600 elicited by number or gender violations individually, we did observe a stronger P600 response in the person+gender violation condition, which we suggest is a consequence of the violation of the person feature. There is much evidence from linguistic and psycholinguistic studies to corroborate the special status of the person feature.

There are a number of possible sources for the special status of person feature violations. First, the phonological contrast between the correct condition and a person violation is greater than violations of other agreement features. It is also possible that orthographic representation may also have had an effect. As can be seen in Table 2 above, within the paradigm of certain verbs the person form is orthographically the most marked. However, this is not clearly the case for all verbs, since the orthographic marking of person agreement varies across different verbs, due to the syllabic nature of *Devanagari* script (see Appendix A for examples of different verb stems and their representations). In other words, the person feature violation may have a special status at the level of surface forms. A second possibility is that the larger P600 response reflects a greater degree of cognitive salience accorded to the person feature, i.e., the person feature violation may have a special status at the level of semantic representation.

A third possibility is that the person feature has a privileged status at the level of morphosyntax, as suggested by converging evidence from language processing and language typology. Based on results from a series of self-paced reading studies Carminati (2005) argues that Italian speakers assign special importance to person agreement information in resolving the reference of an ambiguous null subject. More specifically, Carminati argues that Italian speakers are sensitive to Greenberg's (1963) universal feature hierarchy (person > number > gender), which was originally proposed based on typological analysis of agreement morphology and pronominal systems. Greenberg

proposed that languages reflect this hierarchy in their agreement morphology, in the respect that languages do not encode features that are lower on the hierarchy without also encoding features higher on the hierarchy. Relatedly, it has beeen found that even the simplest pronominal systems reserve unique forms for person while doing away with number and gender distinctions (e.g., the Pirahã language of the Amazon, Thomason and Everett, 2001). Person has also been argued to have a special status in the licensing of null arguments. It has long been known that languages with rich agreement morphology are more likely to allow null arguments, but it has been difficult to specify exactly what counts as sufficiently 'rich' agreement. Vikner (1997) argues that for the Germanic languages the property of verb morphology that best predicts the availability of null arguments is the maintenance of person distinctions in an agreement paradigm. The presence of gender and number distinctions adds to the richness of a paradigm, but does not predict the availability of null arguments. Thus, independent evidence from a number of different sources points to the importance of person as a privileged agreement feature, which may explain the especially large P600 elicited by the person+number violation.

4.3 Lack of LAN Effect

Although a number of previous ERP studies of morphosyntax have shown that morphosyntactic volations often elicit a left anterior negativity (LAN) component (e.g., Coulson et al., 1998; Hagoort & Brown, 2000), no LAN was elicited by the agreement violations in the current study. This is not unprecedented: a number of other studies of agreement violations have observed only a P600 response (Gunter & Friederici, 1999; Hagoort et al., 1993; Lau, Stroud, Plesch, & Phillips, in press; Osterhout & Nicol, 1999). There have even been closely related studies in the same language that have obtained different results in this area: one study of Dutch verb agreement violations (Gunter, 1997) observed a LAN effect that was elicited regardless of the amount of material that intervened between the subject and the verb. In contrast, another study on Dutch agreement violations obtained a LAN effect only when there was little intervening

material (Vos, Gunter, Kolk, & Mulder, 2001). Thus, the factors that govern the appearance of a LAN in response to an agreement violation remain elusive.

5. Conclusion

The aim of the current study was to determine whether the parser is sensitive to the degree of overlap between expected and unexpected agreement marking on a verb. Previous studies on the processing of semantic and acoustic anomalies have demonstrated a feature overlap advantage of this kind, and this effect has been attributed to a priming mechanism. The current study capitalized on the rich verb agreement morphology of Hindi to test whether feature overlap reduces the size of the P600 elicited by an agreement violation. In varying the degree to which an anomalous verb form diverged from the expected form, we found that the parser did not show the same sensitivity to 'goodness-of-fit' that has been observed in other domains. This may reflect qualitative differences between the processing of syntactic and semantic anomalies, or it may reflect differences in the use of predictive mechanisms between syntax and semantics, or it may reflect specific details of the Hindi constructions tested in this study. Further studies of agreement processing languages with rich agreement morphology will be needed in order to resolve these issues. The results also suggested that a larger P600 response is associated with violations of the person feature than violations of other agreement features, a result that is consistent with a variety of previous linguistic and psycholinguistic results that have shown that person has a privileged status among agreement features.

Appendix A: Hindi Verb Agreement

This appendix describes how Hindi future tense verb agreement maps onto *Devanagari* script. Although future tense verb agreement follows a singular regular paradigm, certain aspects of the agreement paradigm are realized in multple ways, due to the fact that *Devanagari* is a syllabic script. This is particularly relevant to the finding in our study that agreement violations involving the person feature elicited a larger P600.

Hindi future tense verb agreement is phonologically regular. Future tense verb forms consist of a verb stem and two suffixes. The first suffix carries information about person and number, and the second suffix carries information about gender and number. These suffixes apply to all verb stems, and the suffixes used in our ERP study are shown in Table 5.

Person/Number		Gender/Number	
-uun-	1st Person, Singular	*-gaa*	Masculine, Singular
-ee-	3rd Person, Singular	*-gee*	Masculine, Plural
-een-	3rd Person, Plural	*-gii*	Feminine, Sing/Plur

Table 5: The phonological forms of the agreement morphemes used in the experiment

Devanagari script is a syllabary, in which each character corresponds to a syllable. Consequently, the realization of each suffix depends on the syllable that the suffix is a part of. Sample forms for three different verb stems used in our ERP study are shown in Table 6. The gender/number suffix includes a syllable onset, and therefore each suffix is consistently realized with the same *Devanagari* character. In contrast, the person/number suffix does not include a consonant as a syllable onset, and therefore this suffix forms a syllable together with the final consonant of the verb stem (or with *–y-* if the stem ends in a vowel). Consequently, each person/number suffix may be realized by a different *Devanagari* character, according to the consonant that the suffix combines with to form a

syllable. This has a potential impact upon the orthographic salience of different agreement violations. As shown in Table 5, 3rd person singular and 3rd person plural suffixes have the same vowel quality and differ only in the presence of nasalization, which is marked by a diacritic mark above the character (see the dot in the 1st person singular and 3rd person plural forms in Table 6). In contrast, 1st person singular and 3rd person singular forms contain different vowels (*-uun-* vs. *–ee-*, respectively), and therefore are realized by a different base *Devanagari* character, according to the syllable onset provided by the final consonant of the stem. There is a good deal of variation in the orthographic salience of the 1st person vs. 3rd person singular contrast, depending on the verb stem. For example, in Table 6 the difference between the suffixes *-uun-* and *-ee-* maps to a large orthographic contrast when they are attached to the stem *gaa* 'sing' and marked by the *Devanagari* characters for *yuun* and *yee*. On the other hand, when the same suffixes appear in the syllables *nuun* and *nee* (*maan* 'obey') or *tuun* and *tee* (*jiit* 'win') the orthographic contrast is less salient. Thus, agreement violations involving the person feature are more salient than other violations when applied to some verb stems, but not when applied to other verb stems.

-uun- (1st, Singular)	*-ee-* (3rd, Singular)	*-een-* (3rd, Plural)
मानूंगी (maanuungii)	मानेगी (maaneegii)	मानेंगी (maaneengii)
गाऊंगा (gaayuungaa)	गायेगा (gaayeegaa)	गायेंगे (gaayeengee)
जीतूंगा (jiituungaa)	जीतेगा(jiiteegaa)	जीतेंगे (jiiteengee)

Table 6: The orthographic forms for three verb stems: *jiit* 'to win', *gaa* 'to sing', and *maan* 'to obey'. The first row of the table contains feminine forms, which end in *–gii*, and the other two rows contain masculine forms, which end in *–gaa* and *–gee*.

References

H. Barber, M. Carreiras, Grammatical Gender and Number Agreement in Spanish: An ERP Comparison, Journal of Cognitive Neuroscience 17 (2005) 135-153.

M.N. Carminati, Processing reflexes of the Feature Hierarchy (Person > Number > Gender) and implications for linguistic theory, Lingua 115 (2005) 259-285.

N. Chomsky. The Minimalist Program, MIT Press, Cambridge, MA. 1995.

S. Coulson, J.W. King, M. Kutas, Expect the unexpected: Event-related brain response to morphosyntactic violations, Language and Cognitive Processes 13 (1998) 21-58.

P. Forchheimer, The Category of Person in Language, Walter de Gruyter, Berlin, 1953.

K.D. Federmeier, M. Kutas, Right words and left words: electrophysiological evidence for hemispheric differences in meaning processing, Cognitive Brain Research 8 (1999) 373-392.

A.D. Friederici, E. Pfeifer, A. Hahne, Event-related brain potentials during natural speech processing: Effects of semantic, morphological, and syntactic violations, Cognitive Brain Research 1 (1993) 183-192.

S. Greenhouse, S. Geisser, On methods in the analysis of profile data, Psychometrika 24 (1959) 95-112.

T. Gunter, A.D. Friederici, Concerning the automaticity of syntactic processing, Psychophysiology 36 (1999) 126-137.

T. Gunter, A.D. Friederici, H. Schriefers, Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction, Journal of Cognitive Neuroscience 12 (2000) 556-568.

T. Gunter, L.A. Stowe, G. Mulder, When syntax meets semantics, Psychophysiology 34 (1997) 660-676.

P. Hagoort, C.M. Brown, Gender electrified: ERP evidence on the syntactic nature of gender processing, Journal of Psycholinguistic Research 28 (1999) 715-728.

P. Hagoort, C.M. Brown, ERP effects of listening to speech compared to reading: the P600/SPS to syntactic violations in spoken sequences and rapid serial visual presentation, Neuropsychologia 38 (2000) 1531-1549.

P. Hagoort, C.M. Brown, J. Groothusen, The Syntactic Positive Shift (SPS) as an ERP measure of syntactic processing, Language and Cognitive Processes 8 (1993) 439-484.

P. Hagoort, M. Wassenaar, C. Brown, Syntax-related ERP-effects in Dutch, Cognitive Brain Research 16 (2003) 38-50.

A. Hahne. What's different in second language processing? Evidence from event-related brain potentials. Journal of Psycholinguistic Research 30 (2001) 251-266.

H. Harley, E. Ritter, Person and number in pronouns: A feature-geometric analysis, Language 78 (2002) 482-526.

R. Jackendoff. Foundations of Language, Oxford University Press, 2002.

E. Kaan, Investigating the effects of distance and number interference in processing subject-verb dependencies: An ERP study, Journal of Psycholinguistic Research 31 (2002) 165-193.

M. Kutas, K.D. Federmeier, Electrophysiology reveals semantic memory use in language comprehension, Trends in Cognitive Sciences 4 (2004) 463-470.

M. Kutas, S.A. Hillyard, Reading senseless sentences: Brain potentials reflect semantic anomaly, Science 207 (1980) 203-205.
M. Kutas, S.A. Hillyard, Event-related brain potentials to grammatical errors and semantic anomalies, Memory & Cognition 11 (1983) 539-550.
M. Kutas, S.A. Hillyard, Brain potentials during reading reflect word expectancy and semantic association, Nature 307 (1984) 161-163.
E.F. Lau, C. Stroud, S. Plesch, C. Phillips, The role of structural prediction in rapid syntactic analysis, Brain and Language (in press).
S. Levänen, R. Hari, L. McEvoy, and M. Sams. Responses of the Human Auditory Cortex to Changes in one versus two Stimulus Features. Experimental Brain Research 97 (1993) 177-183.
T.F. Münte, M. Matzke, S. Johannes, Brain activity associated with syntactic incongruencies in words and pseudo-words, Journal of Cognitive Neuroscience 9 (1997) 318-329.
T.F. Münte, A. Szenkuti, B.M Wieringa, M. Matzke & S. Johannes. Human brain potentials to reading syntactic errors in sentences of different complexity, Neuroscience Letters, 235 (1997) 105–108.
R. Näätänen, Attention and brain function, Erlbaum, Hillsdale, NJ, 1992.
R. Näätänen, A.W.K. Gaillard, S. Mäntysalo, Early selective attention effect on evoked potential reinterpreted, Acta Psychologica 42 (1978) 313–329.
R. Näätänen, M. Tervaniemi, E. Sussman, P. Paavilainen, I. Winkler, 'Primitive intelligence' in the auditory cortex, Trends in Neurosciences 24 (2001) 283-288.
H. Neville, J. Nicol, A. Barss, K.I. Forster, M.I. Garrett, Syntactically-based sentence processing classes: Evidence from event-related brain potentials, Journal of Cognitive Neuroscience 3 (1991) 151-165.
R.C. Oldfield, The assessment and analysis of handedness: the Edinburgh Inventory, Neuropsychologia 9 (1971) 97–113.
L. Osterhout, P.J. Holcomb, D.A. Swinney, Brain potentials elicited by garden path sentences: evidence of the application of verb information during parsing, Journal of Experimental Psychology: Learning, Memory, & Cognition 20 (1994) 768-803.
L. Osterhout, L.A. Mobley, Event-related brain potentials elicited by failure to agree, Journal of Memory and Language 34 (1995) 739-773.
L. Osterhout, Hagoort, P. A superficial resemblance does not necessarily mean you are part of the family: Counterarguments to Coulson, King, and Kutas (1998) in the P600/SPS-P300 debate. Language and Cognitive Processes 14 (1999) 1-14.
L. Osterhout, J. Nicol, On the distinctiveness, independence, and time course of the brain responses to syntactic and semantic anomalies, Language and Cognitive Processes 14 (1999) 283-317.
P. Paavilainen, S. Valppu, and R. Naatenen. The additivity of the auditory feature analysis as indexed by the mismatch negativity: 1+1 = approximately 2, but 1+1+1 < 3, Neuroscience Letters 301 (2001) 179-182.
A.D. Patel, E. Gibson, J. Ratner, M. Besson, Processing syntactic relations in language and music: An event-related potential study, Journal of Cognitive Neuroscience 10 (1998) 717-733.
S. Thomason and D. Everett. Pronoun Borrowing, Proceedings of the Berkeley Linguistics Society 27, 2001.

D.A. Titone, D.F. Salisbury. Contextual modulation of N400 amplitude to lexically ambiguous words, Brain and Cognition 55 (2004) 470–478
C. Van Petten, M. Kutas. Interactions between sentence context and word frequency in event-related brain potentials, Memory and Cognition 18 (1990) 380-393.
C. Van Petten, M. Kutas. Influences of semantic and syntactic context on open and closed class words, Memory and Cognition 19 (1991) 95-112.
S. Vikner. V-to-I Movement and Inflection for Person in All Tenses. In: L. Haegeman, ed., The New Comparative Syntax. Longman, London, (1997). pp. 189-213.
S, Vos, T.C., Gunter, H. Kolk, & G. Mulder. Working memory constraints on syntactic processing: an electrophysiological investigation. Psychophysiology 30 (2001) 41-63.
E. Williams, Remarks on lexical knowledge, Lingua 92 (1994) 7-34.
D. Wunderlich, Minimalist morphology: the role of paradigms. In: G. Booij, J. van Marle (Eds.), Yearbook of Morphology 1995*,* Kluwer*,* 1996, pp. 93-114.