

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/93642>

Please be advised that this information was generated on 2020-09-18 and may be subject to change.

When One Person's Mistake Is Another's Standard Usage: The Effect of Foreign Accent on Syntactic Processing

Adriana Hanulíková^{1,2}, Petra M. van Alphen¹, Merel M. van Goch³,
and Andrea Weber^{1,4}

Abstract

■ How do native listeners process grammatical errors that are frequent in non-native speech? We investigated whether the neural correlates of syntactic processing are modulated by speaker identity. ERPs to gender agreement errors in sentences spoken by a native speaker were compared with the same errors spoken by a non-native speaker. In line with previous research, gender violations in native speech resulted in a P600 effect (larger P600 for violations in comparison with correct sentences), but when

the same violations were produced by the non-native speaker with a foreign accent, no P600 effect was observed. Control sentences with semantic violations elicited comparable N400 effects for both the native and the non-native speaker, confirming no general integration problem in foreign-accented speech. The results demonstrate that the P600 is modulated by speaker identity, extending our knowledge about the role of speaker's characteristics on neural correlates of speech processing. ■

INTRODUCTION

Modern trends of global mobility have increased the number of people who need to communicate in languages other than their mother tongue. As a result, people are encountering non-native (L2) speakers of their language (Cheng, 1999). These L2 speakers are often recognizable by an accent in pronunciation or by grammatical errors, as it is difficult to achieve native-like proficiency (Birdsong & Molis, 2001; Flege, 1995). Despite the differences encountered in language usage, communication normally succeeds, as native (L1) listeners can rapidly adjust to foreign-accented speech and improve their comprehension of it (Bradlow & Bent, 2008; Clarke & Garrett, 2004). They can even become desensitized to L2 grammatical errors. Analyses of natural conversations between L1 and L2 speakers show that L1 speakers rarely correct L2 syntactic errors (Chun, Day, Chenoweth, & Luppescu, 1982) and possibly do not detect otherwise salient mistakes. Given L1 listeners' experience with frequent or fossilized L2 errors, the impact of these mistakes on language comprehension could be modulated by language usage. However, little is known about how and when the brain processes such grammatical mistakes. Using EEG, this study examines the neural consequences of foreign-accented speech

on processing grammatical errors frequently produced by L2 speakers.

Listening to spoken language is usually effortless, despite the fact that listeners have only a very short time to process and integrate various sources of information, including phonology, semantics, syntax, and pragmatics. This proficiency stems from a lifetime of experience with their native language that gives L1 listeners certain expectations about these different sources of information. Such expectations can result in (rather unconscious) anticipations, helping listeners interpret and predict upcoming events (for an overview on mechanisms of expectations or predictions across various domains, see, e.g., Bar, 2007). For example, world knowledge about Dutch trains helps comprehenders interpret the phrase *Dutch trains are white* as false, because Dutch trains are yellow (e.g., Van Berkum, Van den Brink, Tesink, Kos, & Hagoort, 2008; Hagoort, Hald, Bastiaansen, & Petersson, 2004). Similarly, listeners know that a man is unlikely to say *I'm pregnant*, because it is biologically implausible (Van Berkum et al., 2008). Listeners anticipate what might be said and use their world knowledge or stereotype-driven inferences about the speaker, contributing to effective language processing. Studies on changes in the electrical activity of the brain show that conflicts with inferences about what a given speaker might say lead to qualitatively distinct brain ERPs in L1 comprehenders (Van Berkum et al., 2008). Van Berkum et al. presented participants with utterances that could be interpreted as either consistent or inconsistent with the speakers' age, the gender, or the social economic status, as conveyed by the speakers' voice. Van Berkum et al. found that speaker inconsistency (e.g., hearing a man producing the word

¹Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands, ²Basque Center for Cognition, Brain, and Language, Donostia, Spain, ³Radboud University, Nijmegen, the Netherlands, ⁴Donders Institute for Brain, Cognition, and Behaviour, Nijmegen, the Netherlands

pregnant in the sentence *I might be pregnant because I feel sick*) elicited a larger N400 compared with speaker consistency (e.g., hearing a woman producing the word *pregnant* in the same sentence). This modulation of the N400 effect suggests that listeners used what they inferred about the speakers based on their voice characteristics in the earliest stages of meaning construction. It is less clear, however, whether inferences about the speaker could affect neural correlates of syntactic processing.

Listeners can no doubt anticipate likely syntactic structures of their language. For example, listeners expect a noun phrase after the verb *selected* in the English utterance *the scientist selected *to win the prize*. If their expectations are not met, distinctive ERPs to such phrases can be observed (Osterhout & Holcomb, 1992, 1993). Osterhout and Holcomb observed a larger P600 effect to the word *to* in the above low-probability syntactic construction as compared with the same word in a more probable construction such as *the scientist hoped to win the prize*. The modulation of the P600 effect in their study suggests that listeners use the knowledge about probable structures in their L1 during syntactic processing. However, none of the previous studies on syntactic processing varied speaker's characteristics. It is, therefore, less clear whether neural correlates of syntactic processing are also affected by listener's knowledge about frequent or infrequent syntactic structures as a function of speaker's identity.

The majority of studies on syntactic processing employing grammatical conflicts were conducted in the visual modality. In the studies using spoken language, grammatical conflicts were usually produced by L1 speakers (for a review, see Kutas & Federmeier, 2007). However, it can be assumed that L1 comprehenders know that such grammatical errors are uncommon among L1 speakers, whereas they are frequent in L2 speech. This raises the following questions: How do L1 listeners solve syntactic conflicts in L2 speech, which is a more natural environment for error occurrences? And does speaker identity influence the neural dynamics of processing grammatical errors? To answer these questions, we take advantage of the presence of grammatical errors in L2 speakers of Dutch.

Among the most troublesome features of acquiring an L2 is grammatical gender, yet it plays an important role in language processing in many languages. Dutch nouns, for example, belong to either neuter gender, associated with the determiner *het*, or common gender, associated with the determiner *de*. A noun's gender is relevant in adjectival agreement (Booij, 2002). Prenominal adjectives in singular indefinite noun phrases have the inflection *-e* (creating an additional syllable) for common nouns (*groene winkel*, "green store") but not for neuter nouns (*groen huis*, "green house"). Because the gender assignment to nouns differs cross-linguistically (Corbett, 1991), it is often difficult to predict the gender class of a noun in a second language. Unsurprisingly, L2 learners across all proficiency levels and across different languages experience difficulties with gender agreement (Sabourin,

2006; Franceschina, 2005; Guillelmon & Grosjean, 2001), and they have even more difficulties when their L1 lacks grammatical gender (Sabourin, 2006; Franceschina, 2005).

Listeners are sensitive to the correct use of grammatical gender in their L1, as is reflected in studies showing that matching gender is easier to process than mismatching gender (Boelte & Connine, 2004; for a review, see Friederici & Jacobsen, 1999). ERP studies employing contrasts between grammatical violations and correct sentences have shown several components responsive to morphosyntactic agreement involving gender (e.g., Barber & Carreiras, 2005; Gunter, Friederici, & Schriefers, 2000; Münte & Heinze, 1994), such as a LAN and a late posterior positivity (P600). ERP studies in Dutch employing gender violations such as noun phrases preceded by an incorrect article (**het winkel*, "store") or by an incorrectly inflected adjective (**groen winkel*, "green store") have revealed a clear P600 effect (but no LAN) in L1 comprehenders (Van Berkum, Brown, Zwitterlood, Kooijman, & Hagoort, 2005; Hagoort & Brown, 1999). This late positive deflection is assumed to index, among other, processes of syntactic reanalysis or recovery from well-formedness conflicts (Bornkessel-Schlesewsky & Schlewsky, 2009; Friederici, 1995; Hagoort, Brown, & Groothusen, 1993; for a review, see Kutas & Federmeier, 2007; Osterhout & Holcomb, 1992) and may reflect controlled and strategic processes (Hahne & Friederici, 1999; Coulson, King, & Kutas, 1998; Gunter, Stowe, & Mulder, 1997).

If the underlying processes associated with syntactic analysis differ as a function of speaker identity, this should be reflected in the late positive deflection. One of the largest immigrant communities in the Netherlands are Turkish speakers (*Statistics Netherlands*, 2010), thus making many Dutch listeners familiar with Turkish-accented Dutch. Turkish lacks grammatical gender, and corpus studies of L2 Dutch show that Turkish learners of Dutch frequently omit gender marking or use incorrect gender agreement up to 67% of the time (e.g., Orgassa, 2009; Blom, Polišenská, & Weerman, 2006; Weerman, Bisschop, & Punt, 2006). Gender agreement errors are common even in later generations of Turkish immigrants (Cornips, 2008) and are also used as a stereotype about L2 speakers or immigrants (e.g., on Dutch entertainment television programs). In a communicative situation involving L1 speakers, L1 listeners' default expectation is for gender agreement to be well formed (as determined by their linguistic experience). In an interaction with an L2 speaker, in contrast, listeners' default expectations might be governed by higher probabilities of gender errors in foreign-accented speech. This could result in a different strategy of dealing with these mistakes and possibly relaxing their sensitivity to grammatical errors. Previous research indeed suggests that L1 listeners can develop a keen ear for meaning in foreign-accented speech (Cheng, 1999; Galloway, 1980), while focusing less on errors and foreign accents. Listeners could then allocate processing resources away from incorrect syntax in L2 speech. This would not necessarily apply in L1

speech, where the probability of gender errors is considerably lower than in L2 speech. As a consequence, frequent gender agreement errors produced with a Turkish accent might not elicit the same brain responses as would such errors spoken with an L1 Dutch accent.

We used a 2×2 factorial design and compared Dutch listeners' ERP responses to Dutch nouns agreeing or disagreeing in gender with previous context in sentences spoken either by a native speaker of Dutch or by a Turkish learner of Dutch. On the basis of previous research, we predicted that gender violations spoken by the L1 speaker should result in a P600 compared with correct sentences. If listeners are sensitive to higher error probabilities in L2 speech, the same grammatical errors produced by the L2 speaker, however, should result in a modulation of the P600 amplitude. On the basis of previous research on the role of speaker's identity, we predicted either a smaller P600 or no P600 compared with correct sentences. We thus expected to observe an interaction of the two P600 effects as a function of the accent of the speaker.

To control for possible effects of shallow processing or integration difficulties in foreign-accented speech compared with native speech, a control set of sentences including semantic world knowledge violations was also created. An ERP response frequently observed for these violations is the N400 effect (e.g., Kutas & Federmeier, 2007; Kutas & Hillyard, 1980). A consistent finding in the literature on the N400 effect is that its amplitude is negatively correlated with the fit of a word in the (semantic) context. We hypothesized that semantic violations such as *I put a thick evening on my bed* create a conflict with both L1 and L2 speakers' typical productions, because neither of the two speaker groups is likely to produce such sentences. We therefore predicted that these violations should elicit comparable N400 effects in both native and foreign-accented speech. This would confirm that there are no general comprehension difficulties because of a foreign accent.

METHODS

Participants

Thirty-four native speakers of Dutch with no hearing, neurological, or psychiatric disorders volunteered to participate. They were students (17 women, all right handed except for two ambidextrous, mean age = 22 years, range = 18–30 years) at Radboud University Nijmegen. None of the students reported a Turkish background or knowledge of the Turkish language. After the experiment, 28 participants correctly recognized that the speaker was from Turkey. One of the participants reported not to be familiar with the foreign accent of the speaker, two participants reported the speaker to be Moroccan, and the remaining three participants reported hearing a German, an Iraqi, and a Japanese accent, respectively.

Materials and Pretests of Materials

The critical stimuli consisted of 240 sentences, selected from an initial set of 320 sentences after auditory recordings and pretesting (see below). Each sentence was recorded in four versions; a correct and an incorrect version of each sentence were spoken by a native speaker and by a non-native speaker (resulting in 960 sentences). Half of the incorrect sentences contained gender disagreement between the definite determiner and the noun, and the other half contained incorrectly inflected adjectives (see Table 1 for example sentences). A further set of 104 control sentences was selected from 120 sentences after auditory recordings. Each sentence was recorded in four versions in the same way as the critical stimuli (resulting in 416 sentences). The incorrect sentences contained semantic violations (see Table 1). The critical nouns at which the grammatical or semantic violation became apparent were embedded in a subordinate clause following a main clause, at least five syllables before the end of the subordinate clause. All critical nouns were balanced for gender (half neuter) and frequency (using the Spoken Dutch Corpus; Oostdijk, 2000).

The sentences were spoken by a female Dutch native speaker and by a female Turkish speaker of Dutch, respectively. The Turkish speaker spoke Dutch fluently, but with a clear foreign accent, as determined with accent ratings stemming from a questionnaire collected in a pretest (see below). On a scale from 1 (*strong foreign accent*) to 5 (*no foreign accent*), the Turkish speaker had an average of 2.4. Both speakers received a list of all sentences in both versions (correct and incorrect) in a randomized order. They were asked to read the sentences at a natural speech rate. To minimize possible differences in the speech rate and the intonation across speakers and conditions, each sentence was first produced by the native Dutch speaker

Table 1. Sentences with English Translation

Sentences with Gender Agreement

Ik wil een reis naar China maken, omdat de/*het *cultuur* daar zo anders is dan hier.

"I want to make a trip to China, because the *culture* there differs from the one over here."

Mijn moeder belde in paniek op, omdat een duur/*dure *juweel* uit haar tas was gestolen.

"My mother panicked because expensive *jewelry* had been stolen from her bag."

Control Sentences with Semantic Manipulation

Het was vannacht best koud, dus ik had een dikke *deken*/**avond* op mijn bed gelegd.

"It was very cold last night, so I put a thick *blanket/evening* on my bed."

Critical words are in italics. Asterisk indicates an incorrect word in a given context.

and immediately repeated by the Turkish speaker of Dutch.¹ In native speech, the mean duration of the critical words was 411 msec ($SD = 98$ msec) and the mean duration of the whole sentence was 4993 msec ($SD = 582$ msec). In non-native speech, the mean duration of the critical word was 387 msec ($SD = 101$ msec) and 5124 msec ($SD = 614$ msec) of the whole sentence.² All sentences were adjusted for amplitude. The mean logarithmic critical word form frequency per million was 2.23 ($SD = 0.61$) in sentences with the gender agreement manipulation and 1.93 ($SD = 0.63$) in sentences with the semantic manipulation.

To select the best material, all recorded sentences containing gender agreement manipulations were pretested in an off-line error detection task. Four randomized lists were created with only one of the four different versions of each sentence occurring in a given list. Participants (20 L1 Dutch university students in Nijmegen, mean age = 21 years, range = 18–25 years, 18 women, none of whom participated in the EEG experiment) were asked to press a response button when they heard a grammatical error. No information about the nature of the errors was provided. Sentences in which errors were either missed or false-reported by more than five participants were excluded from the EEG experiment. Participants showed 93% correct error detections in non-native speech and 97% correct error detections in native speech (one participant was excluded because of a misdetection rate of 3 SD below the participants' average).³

Four experimental lists were then created with one version of each of the 240 critical sentences and the 104 control sentences occurring only in one experimental list (each list thus contained altogether 344 sentences). The frequency of the critical words was matched across conditions. In total, 35% ($n = 120$) of all sentences within an experimental list contained a gender violation (half of which were spoken by the native speaker, and the other half were spoken by the Turkish speaker), and 15% ($n = 52$) of all sentences contained a semantic violation (half of which were spoken by the native speaker, and the other half were spoken by the Turkish speaker).

Procedure

Participants were seated in a comfortable chair and were told that they would hear two speakers talking about their lives. No information about the background or the accent of the speakers was provided. The 344 utterances were presented over loudspeakers. Participants were asked to listen for comprehension and were told that they would have to answer comprehension questions following some sentences. The questions (12 yes or no questions, half of which required a “yes” response) were added to ensure that participants paid attention during the experiment. No grammaticality judgment or acceptability task was used.⁴ After the presentation of each utterance, a cross appeared in the middle of the screen to indicate that participants

could blink or move. Participants were given button-press control over the initiation of the next trial, which started with a silence of 1000 msec followed by the utterance. The experiment consisted of eight blocks (each block lasted on average 8 min) and seven breaks. Participants completed a language-background questionnaire after the EEG study.

EEG Recording and Analysis

EEG was recorded from 34 Ag–AgCl electrodes (impedance was kept below 10 k Ω) at standard 10–20 locations (Fz, Cz, Pz, Oz, Fp1/2, F3/4, F7/8, F9/10, FC1/2, FC5/6, FT9/10, C3/4, T7/8, CP1/2, CP5/6, P3/4, P7/8, O1/2, PO9/10). All recordings were referenced to the left mastoid during recording (eye movement and blink artifacts were recorded from F9 to F10 and from Fp1 to an additional EOG electrode below the left eye), amplified with BrainAmp DC amplifiers (0.016–100 Hz band pass, digitized at 500 Hz), and re-referenced off-line to the mastoid average. EEG segments ranging from 200 msec before to 1500 msec after critical word onset were extracted and baseline corrected to a 200-msec pre-onset baseline. Two participants were excluded because of recording problems, a further participant was excluded because of extensive alpha (exceeding 3 SD from the mean alpha across all participants and all trials), and one participant was excluded because his error rate on the comprehension questions during the EEG session exceeded 3 SD from participants' average (4%). Across the remaining 30 participants, the segments with potentials above ± 75 μ V were marked as artifacts and deleted from further analyses (average segment loss = 13%, range = 12–15%, no difference between conditions). The segments were averaged per participant and condition, and mean amplitudes in specific time windows were analyzed with repeated measures ANOVAs. First, the variation of effect size over all electrodes was evaluated, after which a topography-oriented analysis was conducted involving anterior (Fp1/2, F3/4, F7/8, FC1/2, FC5/6, F9/10, FT9/10, Fz) and posterior distribution (CP5/6, CP1/2, P7/8, P3/4, PO9/10, O1/2, Pz, Oz).

For the statistical analyses, the time window after the critical word onset was chosen in line with previous research within the auditory modality and on the basis of visual inspection of the averaged data. For the P600 effect, the time window was 800–1200 msec (for a similar time window, see, e.g., Friederici & Oberecker, 2008; Rossi, Gugler, Hahne, & Friederici, 2005; Koester, Gunter, Wagner, & Friederici, 2004), and for the N400 effect, it was 300–500 msec (see, e.g., Kutas & Federmeier, 2007). Note that the detection of the agreement error is only possible once the critical noun has been heard and recognized and the gender of the noun has become available. The resulting P600 effect can therefore be delayed in speech. Although the N400 condition also requires the recognition of the critical noun, the retrieval of the noun's

gender information is not necessary for the detection of the semantic incongruity.

RESULTS

As can be seen in Figure 1A, gender violations in native speech resulted in a larger P600 compared with correct sentences, with a clear posterior distribution. The variation of the effect size over all electrodes was confirmed in an ANOVA with the factors Correctness (violation, control) and Electrodes (all 34), revealing a significant interaction in the 800–1200 msec window ($F(1, 33) = 6.61, p < .001, \eta_p^2 = .186$). To examine whether the effect indeed had a posterior distribution, a topography-oriented analysis was conducted, with the factors Distribution and Correctness. This analysis confirmed that the effect was larger over the posterior than the anterior area (Distribution \times Correctness: $F(1, 29) = 12.21, p = .002, \eta_p^2 = .296$). Follow-up analyses revealed a significant P600 effect to violations compared with correct sentences across all posterior electrodes ($F(1, 29) = 8.66, p = .006, \eta_p^2 = .230$) but not across all anterior electrodes ($F < 1$). No other significant differences were found in an earlier time window of 300–500 msec.

Interestingly, when the same listeners heard the same type of gender violations in non-native speech, no such P600 effect was observed (see Figure 1B). The lack of variation in effect size across all electrodes was confirmed by a nonsignificant interaction of the factors Correctness and Electrodes ($F(1, 33) = 1.95, p = .10, \eta_p^2 = .063$).

None of the topography-oriented and follow-up analyses showed significant effects, and no other significant differences were found in an earlier time window of 300–500 msec.

To examine whether the size of the P600 effect depended on the accent of the speaker, we compared the ERP effect elicited by the native speaker to that elicited by the non-native speaker, considering posterior electrodes only. A 2 (Accent of the Speaker) \times 2 (Correctness) ANOVA confirmed that the two effects differed in size across posterior electrodes ($F(1, 29) = 7.18, p = .01, \eta_p^2 = .198$), suggesting that the P600 effect was modulated by the accent of the speaker.

Although the results so far suggest that the P600 effect was only present in L1 speech, we examined the possibility that the P600 effect was present in L2 speech too but diminished during the course of the experiment because of an adjustment to grammatical errors. For this purpose, the data was split into two blocks corresponding to the first and second halves of the experiment. As can be seen in Figure 2, there was a clear posteriorly distributed P600 to violations compared with correct sentences spoken by the L1 speaker in the first block (as confirmed by a main effect of condition over posterior electrodes, $F(1, 29) = 11.94, p = .002, \eta_p^2 = .292$). This was not found in the second block ($F < 1$). For the L2 speaker, neither of the two blocks showed a P600 effect over posterior electrodes (both F s < 1). Follow-up analyses over all posterior electrodes and both accents confirmed that the size of the P600 effect was modulated by the accent of the

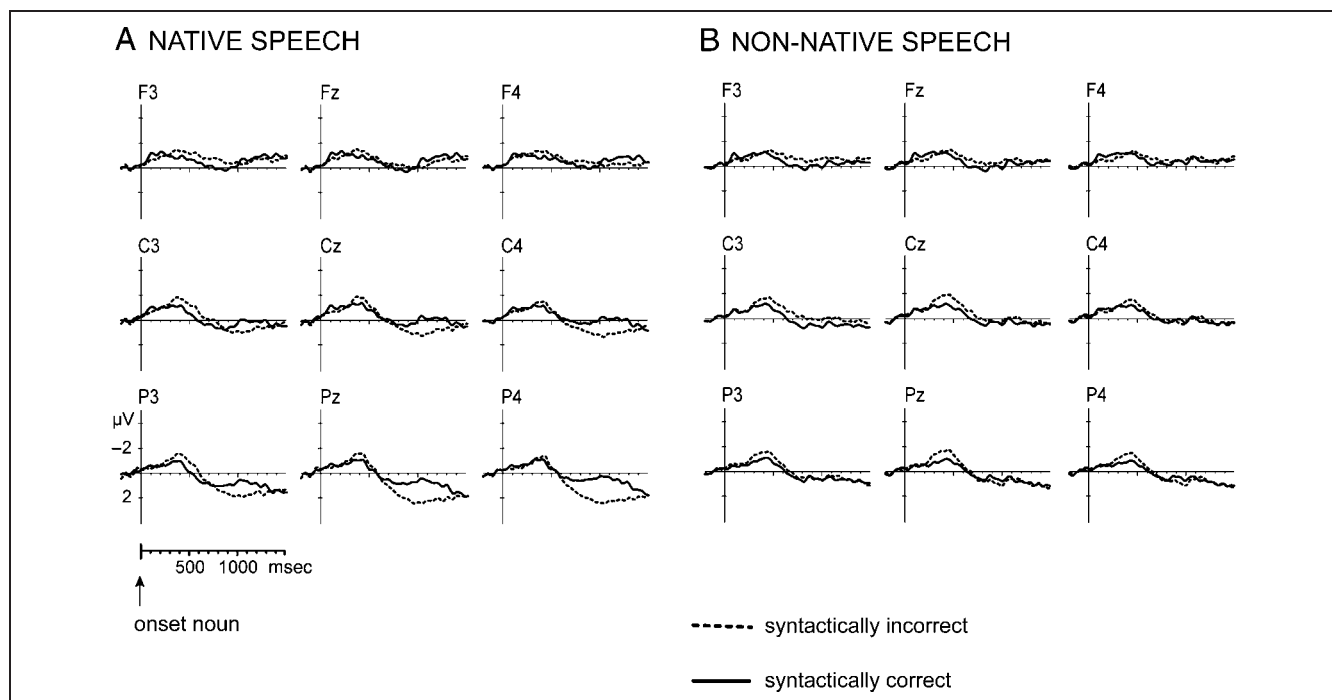
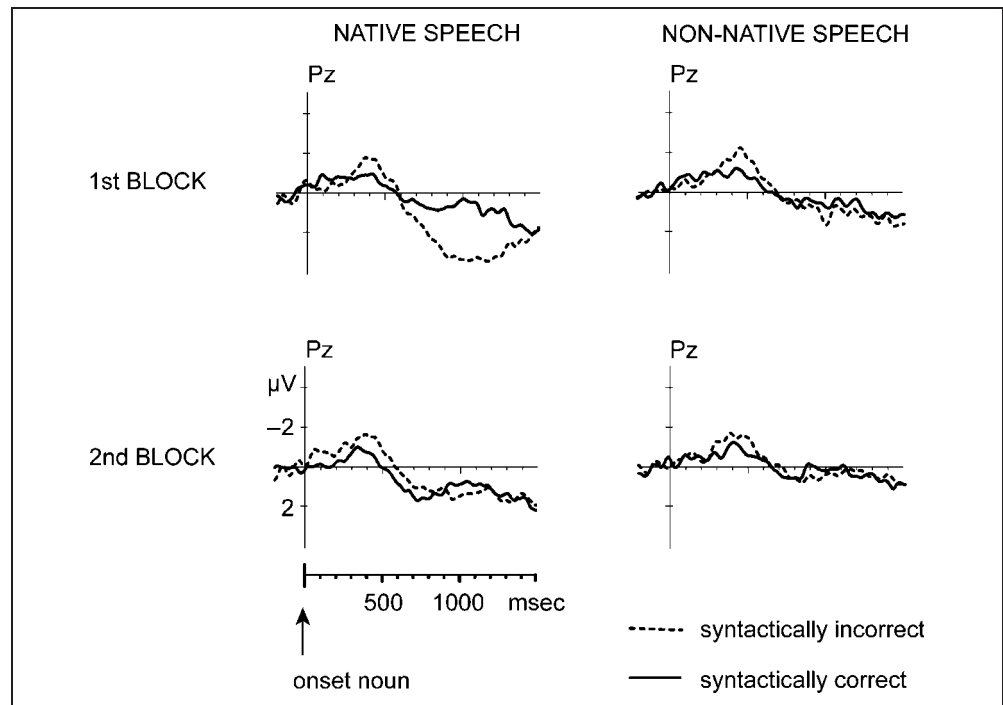


Figure 1. (A, B) Grand average ERPs from nine scalp sites elicited by syntactically incorrect nouns (dashed lines) and syntactically correct nouns (solid lines) in native speech (A) and non-native speech (B). Waveforms are filtered (5 Hz high cutoff, 12 dB/oct) for presentation purpose only.

Figure 2. Grand average ERPs from the parietal site Pz elicited by syntactically incorrect nouns (dashed lines) and syntactically correct nouns (solid lines) in native and non-native speech in the first and second block.



speaker in the first block (Accent of the Speaker \times Correctness: $F(1, 29) = 4.99, p = .03, \eta_p^2 = .147$), but not in the second block ($F < 1$). We return to this result in the discussion below.

To exclude the possibility of shallow processing or integration problems in foreign-accented speech, control sentences with semantic manipulation were analyzed. As

can be seen in Figure 3A and B, semantic violations elicited a larger posteriorly distributed N400 than correct sentences in the 300–500 msec window for both the native speaker and the non-native speaker. An ANOVA with the factors Correctness and Electrodes (34) revealed a significant interaction for both the native speech ($F(1, 33) = 6.54, p < .001, \eta_p^2 = .184$) and the non-native speech ($F(1, 33) = 5.87,$

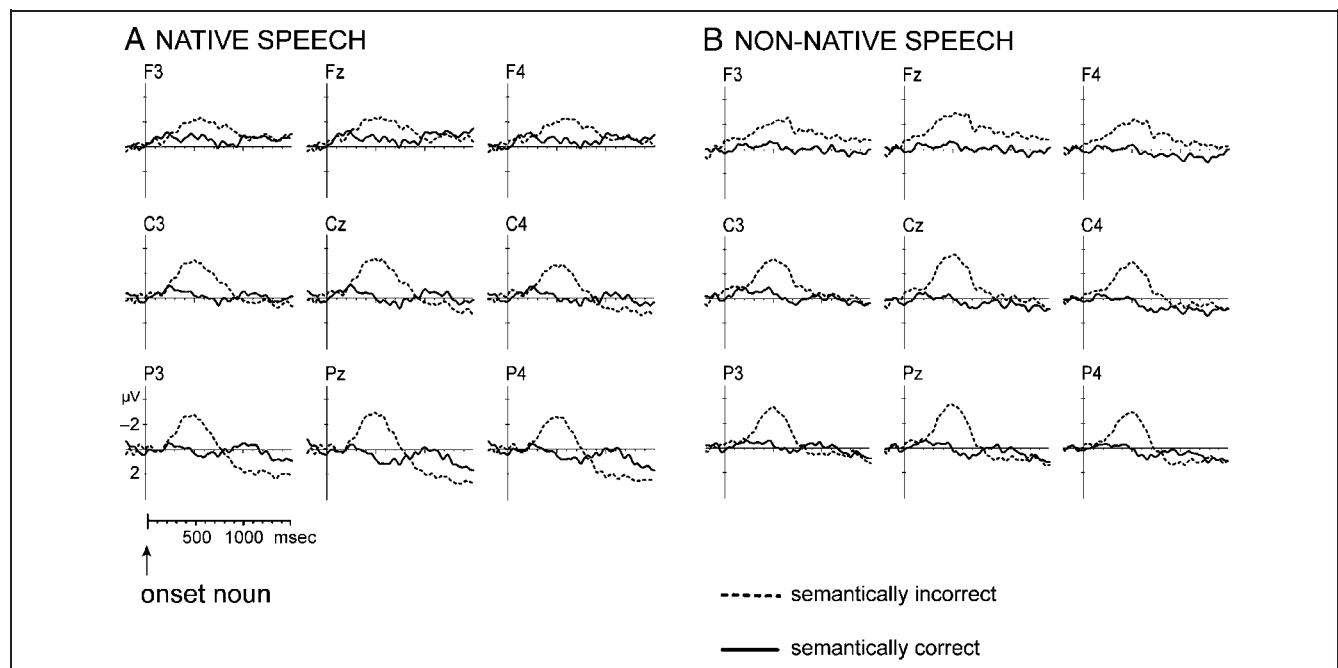


Figure 3. (A, B) Grand average ERPs from nine scalp sites elicited by semantically correct nouns (dashed lines) and semantically incorrect nouns (solid lines) in native speech (A) and non-native speech (B).

$p < .001$, $\eta_p^2 = .168$). These results confirm that the N400 effects varied in size across all electrodes. A topography-oriented analysis showed that the effect was indeed larger over the posterior than the anterior area for the native speech (Distribution \times Correctness: $F(1, 29) = 8.32$, $p = .007$, $\eta_p^2 = .223$) but not for the non-native speech ($F(1, 29) = 1.62$, $p = .21$, $\eta_p^2 = .053$). Follow-up analyses revealed a significant N400 effect across all posterior electrodes for both the native speech ($F(1, 29) = 22.71$, $p < .001$, $\eta_p^2 = .439$) and the non-native speech ($F(1, 29) = 13.54$, $p = .001$, $\eta_p^2 = .318$). Across all anterior electrodes, the N400 was marginally significant in the native speech ($F(1, 29) = 3.73$, $p = .06$, $\eta_p^2 = .114$) and reliable in the non-native speech ($F(1, 29) = 7.61$, $p = .01$, $\eta_p^2 = .208$).

The size of the posterior N400 effect did not depend on the accent of the speaker (Accent of the Speaker \times Correctness: $F < 1$), suggesting that the N400 effect was not modulated by the accent of the speaker. Although the results indicated that the N400 had a broader distribution for the non-native speaker (anterior and posterior distribution) than for the native speaker (posterior distribution only), the size of the anterior N400 effect did not depend on the accent of the speaker (Accent of the Speaker \times Correctness: $F < 1$). This result confirms that listeners were attending equally well to sentences spoken by both the native speaker and the non-native speaker.

DISCUSSION

Unlike native speech, the speech of second language learners typically contains grammatical errors. However, there is a dearth of studies on this more intuitive scenario in which errors are likely to occur in actual speech, despite the large literature on ERP effects observed to grammatical violations. The present study showed that gender agreement errors led to a change in the electrophysiological response to the subsequent noun when L1 listeners heard these mistakes made by an L1 speaker, but not when the same mistakes were made by an L2 speaker with a clear non-native accent. Importantly, L1 listeners did not appear to experience comprehension or shallow processing problems in L2 speech, as indicated by almost equivalent electrophysiological responses to world knowledge violations produced by either the L1 speaker or the L2 speaker. Although processing world knowledge in L2 speech thus did not differ from L1 speech, syntactic processing as reflected by the P600 clearly was affected.

What is the source of this difference in the electrophysiological response to grammatical errors? It seems that listeners' previous experiences with the correlation between L2 accent and error likelihood have modified their expectations about the grammatical well-formedness of foreign-accented speech. Consequently, this could have led to a differential impact of grammatical errors on processing. It can be assumed that L1 comprehenders have experienced that the most problematic aspects of L2 acquisition are morphosyntax and pronunciation, both of which do not

necessarily preclude successful communication in the absence of native-like proficiency. Persevering gender agreement errors, for example, could be relatively well overlooked during the process of sense making, because they might not affect the communicative intention of the speaker. On the basis of the accumulated evidence of such L2 errors in real life, listeners might have developed sensitivity to distributional differences of grammatical errors between L1 and L2 speech and could thus reduce attempts to repair these grammatical errors. The absence of a P600 effect in the present study might reflect such a reduction of grammatical repairs.⁵ In native speech, in contrast, gender errors are less probable. L1 listeners are therefore more likely to attempt grammatical repair processes, as indicated by a P600 in the present study. This would suggest that L1 listeners have adjusted their probability model about error occurrences in L2 speech but not in L1 speech.

An adjustment to errors seems in line with our block analysis showing no P600 effect to grammatical violations spoken by the L2 speaker in either of the two blocks. Interestingly, grammatical violations spoken by the L1 speaker elicited a P600 only in the first but not the second block. This suggests that adjustments to errors occur even in L1 speech. However, unlike in L2 speech, the adjustments in L1 speech seem to be rather gradual. It thus appears that listeners adjust their probability model about error occurrences resulting in fewer attempts of grammatical repairs even when errors are repeatedly produced by an L1 speaker. This is in line with previous studies that showed flexibility of brain responses to L1 grammatical errors as a function of error probability within an experiment (e.g., Hahne & Friederici, 1999; Coulson et al., 1998; Gunter et al., 1997). Hahne and Friederici, for example, compared ERP responses to phrase structure violations with responses to correct sentences when the proportion of violations within an experiment was either 20% or 80% of all trials. L1 comprehenders showed a P600 to violations in the 20% condition, but the P600 was largely diminished in the 80% condition. Although no split analysis was provided to directly evaluate the adjustment to errors across the course of the experiment, the overall results suggest that listeners refrained from repair attempts because of the extremely high proportion of errors even in L1 speech. In other words, if an error becomes the "standard" (as in the high error proportion manipulation), listeners are able to adjust to these errors and can direct their attention to what they hear in a way needed for a successful comprehension. Note that the present study contained a constant number of grammatical violations (35%) in both the L1 and the L2 speech. The modulation of the P600 effect was, therefore, not driven by proportion manipulation within the study but rather by inferences about what a certain speaker is likely to say. How quickly listeners adjust can be, thus, also modulated by speaker's identity. As the present study demonstrates, adjustments in brain responses observed to grammatical violation depend on whether the speaker has a native or a non-native accent.

Why would adjustments differ as a function of the accent of the speaker? One possibility is that the outcome of the block analysis reflects a short-term adjustment to errors during the experiment for the L1 speaker and a long-term adjustment to errors for the L2 speaker. Long-term adjustment results from the amount of accumulated experience with L2 speakers, leading to changes in estimates of the probability of correct formal features. Non-native speech often contains errors. Moreover, it is associated with a salient cue to the different probability distribution—namely, the foreign accent (indeed, over 80% of the participants recognized the Turkish accent). For these reasons, the error patterns in non-native speech might be easier to overlook, whereas the error patterns in native speech are less typical and require a certain amount of evidence to achieve adjustment for a specific native speaker. A second possible interpretation of the adjustment results could be based on a stronger decline in attention for L2 speech during the experiment. In this case, however, a similar decline in attention would be expected for the N400 effect. We examined this possibility with an additional block analysis of the semantic condition, which revealed a constant N400 over blocks in both the L1 and the L2 speech. This suggests that adjustment to syntax but not to world knowledge was modulated by prior experience.

Overall, this result extends our knowledge about the role of speaker's characteristics on neural correlates of speech processing. Previous research has already shown that listeners take into account speaker's identity during meaning construction (Van Berkum et al., 2008). Speaker inconsistency, such as the improbable sentence *I am pregnant* uttered by a male voice elicited a larger N400 than the same sentence uttered in a more probable context of a female voice. The present study has shown that syntactic processing as reflected by the P600 too can be modulated in a similar manner. Inconsistencies, on the basis of infrequent events, such as a native speaker producing a gender error, elicited a larger P600 compared with the same gender errors produced in a more likely context of foreign-accented speech. The linguistic brain thus takes into account all information available to achieve an effortless and successful comprehension of spoken language.

This study not only suggests a certain degree of tolerance with respect to a suboptimal syntactic fit in L2 speech but also shows tolerance with respect to the foreign accent. Many studies on effects of foreign-accented speech on language comprehension (for a review, see Munro & Derwing, 1995) propose that phonology interferes with comprehension to a greater extent than syntax (Van Heuven, 1986). Other studies claim that grammar is more detrimental than phonology (Ensz, 1982), whereas yet others believe that the probability of a given error is a more decisive factor in the comprehension of L2 speech (Albrechtsen, Henriksen, & Færch, 1980). However, a foreign accent does not necessarily impinge upon communication. Ultimately, the listener's concern is to rapidly extract meaning from the speech sig-

nal and, as the present study shows, listeners succeed in doing so relatively well in both native and non-native speech, even if it sometimes requires overlooking the syntactic rules of their language. Future research needs to examine whether other types of mistakes by L2 speakers show similar results. It is possible that not all errors are equal in gravity and that grammatical errors altering meaning affect parsing more than errors which are less crucial for meaning.

The present results also speak to a larger debate about the functional interpretation of the P600 effect. The P600 effect has been often associated with the processing of incorrect syntactic constructions (Hagoort et al., 1993) as well as infrequent syntactic constructions (Osterhout & Holcomb, 1992, 1993). Several studies have shown a dependence of the P600 response on manipulations of participants' expectations concerning the frequency of error occurrences within an experiment (e.g., Hahne & Friederici, 1999; Coulson et al., 1998; Gunter et al., 1997; but see Osterhout, McKinnon, Bersick, & Corey, 1996, who do not confirm such dependency). On the basis of these studies, the P600 has been described as a controlled rather than an automatic response. Some researchers have therefore argued that the P600 is not language specific but that it is part of the P300 family (e.g., Coulson et al., 1998). Specifically, it resembles the P3b component, which is sensitive to probability manipulations of an event. The present result indeed confirms that the late positivity is modulated by error probability, and shows that the error probability can be inferred on the basis of the accent of the speaker as well as on manipulations of the amount of errors within an experiment. Thus, the late positivity can be modulated by participants' inferences about error occurrences as a function of speaker. Although the present results do not allow resolving the debate on whether this late negativity is a P600 or rather part of the P300 family, they do demonstrate that the late component is a suitable tool for investigation of experiential factors in spoken-language processing. Future research could further explore this issue and examine the effect of accent of the speaker on earlier, automatic syntactic processes.

In conclusion, achieving mutual understanding despite the large amount of variation in spoken language is an amazing human achievement. People are rapid in adapting to new situations and new speakers, and are able to reduce interference effects following conflicts. Given that cognitive control is dynamic, monitoring of conflicts leads to behavioral adjustments. Listeners can effectively use a foreign accent as a cue for non-nativeness and adjust their probability model to make the communication a successful enterprise. Although L1 listeners can certainly detect grammatical errors in L2 speech if explicitly required to do so, the impact that these errors have on them in a naturalistic listening situation is not detrimental. Moreover, L1 comprehenders can adjust to real-life probabilities of error occurrence. This is good news for L2 speakers who are often embarrassed for producing grammatical errors.

Overlooking grammatical errors in a conversation could then be seen as a natural and automatic attempt to ensure successful communication between native and non-native speakers.

Acknowledgments

We thank Laurence Bruggeman and Karina Visser for their help with sentence construction and Laurence Bruggeman and Anne Blankenhorn for assistance in testing participants. We also thank three anonymous reviewers for their constructive feedback on the manuscript. This research was funded by the Max-Planck-Gesellschaft, Germany.

Reprint requests should be sent to Adriana Hanulíková, Basque Center on Cognition, Brain and Language, Mikeletegi Pasealekua 69, Solairua 2, 20009 Donostia, Spain, or via e-mail: Adriana.Hanulikova@mpi.nl.

Notes

1. No editing or cross-splicing of the critical nouns was administered to avoid coarticulation inconsistencies across different conditions (e.g., the phonetic properties of the onset of the critical noun are affected by the preceding speech sound, which differed across conditions because of differing determiners or adjectival inflections).
2. The differences in duration between the native and the accented speech were significant for both sentence duration ($t(687) = 15.30, p < .001$) as well as word duration ($t(687) = 9.24, p < .001$).
3. The high error detection rates suggest that the morphological manipulations were phonologically salient. Note that the correct and incorrect adjectives preceding half of the critical words differed in the number of syllables (e.g., *groen* vs. *groene*).
4. One advantage of using EEG is that it allows investigating brain activity as the speech unfolds over time, without an additional task that could interfere with the natural spoken language processing. It is often argued that the P600 is more likely to be elicited with a judgment task. However, numerous task-less studies have elicited ERP effects, some of which are closely related to the present study (e.g., Van Berkum et al., 2008; Hagoort & Brown, 2000). To keep the task as natural as possible and to keep the study comparable to previous studies, we asked participants to listen for comprehension only.
5. One might wonder whether a P600 in L2 speech could be observed with a sentence judgment task. Judging a sentence may strengthen control processes, possibly resulting in a modulation of the ERP effects. However, the aim of the present study was to better understand the natural situation of processing grammatical errors produced by L1 and L2 speakers rather than to investigate under which task-related conditions a P600 effect might be elicited.

REFERENCES

- Albrechtsen, D., Henriksen, B., & Færch, C. (1980). Native speaker reactions to learners' spoken interlanguage. *Language Learning, 30*, 365–396.
- Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences, 11*, 280–289.
- Barber, H., & Carreiras, M. (2005). Grammatical gender and number agreement in Spanish: An ERP comparison. *Journal of Cognitive Neuroscience, 17*, 137–153.
- Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language, 44*, 235–249.
- Blom, E., Polišenská, D., & Weerman, F. (2006). Effects of age on the acquisition of agreement inflection. *Morphology, 16*, 313–336.
- Boelte, J., & Connine, C. M. (2004). Grammatical gender in spoken word recognition in German. *Perception and Psychophysics, 66*, 1018–1032.
- Booij, G. E. (2002). *The morphology of Dutch*. Oxford, UK: Oxford University Press.
- Bornkessel-Schlesewsky, I., & Schlewsky, M. (2009). *Processing syntax and morphology: A neurocognitive perspective*. Oxford, UK: Oxford University Press.
- Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition, 106*, 707–729.
- Cheng, L. R. L. (1999). Moving beyond accent: Social and cultural realities of living with many tongues. *Topics in Language Disorders, 19*, 1–10.
- Chun, A. E., Day, R. R., Chenoweth, N. A., & Luppescu, S. (1982). Errors, interaction, and correction: A study of native-nonnative conversations. *TESOL Quarterly, 16*, 537–547.
- Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *Journal of Acoustical Society of America, 116*, 3647–3658.
- Corbett, G. (1991). *Gender*. Cambridge, MA: Cambridge University Press.
- Cornips, L. (2008). Loosing grammatical gender in Dutch: The result of bilingual acquisition and/or an act of identity? *International Journal of Bilingualism, 12*, 105–124.
- Coulson, S., King, J. W., & Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes, 13*, 21–58.
- Ensz, K. Y. (1982). French attitudes toward typical speech errors of American speakers of French. *The Modern Language Journal, 64*, 210–215.
- Flege, J. E. (1995). Second language speech learning: Theory, findings and problems. In W. Strange (Ed.), *Speech perception and linguistic experience* (pp. 233–272). Baltimore: York Press.
- Franceschina, F. (2005). *Fossilized second language grammars: The acquisition of grammatical gender*. Amsterdam: Benjamins.
- Friederici, A. D. (1995). As time goes by: The time-course of syntactic activation during language processing. *Brain and Cognition, 28*, 259–281.
- Friederici, A. D., & Jacobsen, T. (1999). Processing grammatical gender during language comprehension. *Journal of Psycholinguistic Research, 28*, 467–484.
- Friederici, A. D., & Oberecker, R. (2008). The development of syntactic brain correlates during the first years of life. In A. D. Friederici & G. Thierry (Eds.), *Early language development: Bridging brain and behaviour* (pp. 215–231). Amsterdam: John Benjamins.
- Galloway, V. B. (1980). Perception of the communicative efforts of American students of Spanish. *The Modern Language Journal, 64*, 428–433.
- Guillelmon, D., & Grosjean, F. (2001). The gender marking effect in spoken word recognition: The case of bilinguals. *Memory and Cognition, 29*, 503–511.
- Gunter, T. C., Friederici, A. D., & Schriefers, H. (2000). Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. *Journal of Cognitive Neuroscience, 12*, 556–568.
- Gunter, T. C., Stowe, L. A., & Mulder, G. (1997). When syntax meets semantics. *Psychophysiology, 34*, 660–676.

- Hagoort, P., & Brown, C. M. (1999). Gender electrified: ERP evidence on the syntactic nature of gender processing. *Journal of Psycholinguistic Research*, 28, 715–728.
- Hagoort, P., & Brown, C. M. (2000). ERP effects of listening to speech compared to reading: The P600/SPS to syntactic violations in spoken sentences and rapid serial visual presentation. *Neuropsychologia*, 38, 1531–1549.
- Hagoort, P., Brown, C. M., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes*, 8, 439–483.
- Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, 304, 438–441.
- Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis. Early automatic and late controlled processes. *Journal of Cognitive Neuroscience*, 11, 194–205.
- Koester, D., Gunter, T. C., Wagner, S., & Friederici, A. D. (2004). Morphosyntax, prosody, and linking elements: The auditory processing of German nominal compounds. *Journal of Cognitive Neuroscience*, 16, 1647–1668.
- Kutas, M., & Federmeier, K. D. (2007). Event-related brain potential (ERP) studies of sentence processing. In G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 385–406). Oxford, UK: Oxford University Press.
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207, 203–205.
- Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45, 73–97.
- Münte, T. F., & Heinze, H. J. (1994). ERP negativities during syntactic processing of written words. In H. J. Heinze, T. F. Münte, & G. R. Magnun (Eds.), *Cognitive electrophysiology*. Boston: Birkhäuser.
- Oostdijk, N. (2000). The Spoken Dutch Corpus project. *The ELRA Newsletter*, 5, 4–8.
- Orgassa, A. (2009). *Specific language impairment in a bilingual context: The acquisition of Dutch inflection by Turkish-Dutch learners*. Utrecht: LOT.
- Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, 31, 785–806.
- Osterhout, L., & Holcomb, P. J. (1993). Event-related potentials and syntactic anomaly: Evidence of anomaly detection during the perception of continuous speech. *Language and Cognitive Processes*, 8, 413–437.
- Osterhout, L., McKinnon, R., Bersick, M., & Corey, V. (1996). On the language specificity of the brain response to syntactic anomalies: Is the syntactic positive shift a member of the P300 family? *Journal of Cognitive Neuroscience*, 8, 507–526.
- Rossi, S., Gugler, M. F., Hahne, A., & Friederici, A. D. (2005). When word category information encounters morphosyntax: An ERP study. *Neuroscience Letters*, 384, 228–233.
- Sabourin, L. (2006). Transfer effects in learning a second language grammatical gender system. *Second Language Research*, 22, 1–29.
- Statistics Netherlands. (2010). <http://statline.cbs.nl/StatWeb/publication/?DM=SLEN&PA=37325eng&D1=0&D2=0&D3=0&D4=0&D5=0-1,84,102,139,145,210,225&D6=a&LA=EN&HDR=G2,G3,G4,T&STB=G1,G5&VW=T>. Retrieved on June 1, 2010.
- Van Berkum, J. J., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 443–467.
- Van Berkum, J. J., Van den Brink, D., Tesink, C. M., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience*, 20, 580–591.
- Van Heuven, V. J. (1986). Some acoustic characteristics and perceptual consequences of foreign accent in Dutch spoken by Turkish immigrant workers. In J. Van Oosten & J. F. Snapper (Eds.), *Dutch Linguistics at Berkeley, papers presented at the Dutch Linguistics Colloquium held at the University of California, Berkeley on November 9th, 1985* (pp. 67–84). U.C. Berkeley, CA: Berkeley's Dutch Studies Program.
- Weerman, F., Bisschop, J., & Punt, L. (2006). L1 and L2 acquisition of Dutch adjectival inflection. *ACL Working Papers*, 1, 5–36.