Foreign-accented speaker identity affects neural correlates of language comprehension

Sarah Grey a, b, *, Janet G. van Hell a

a Department of Psychology and Center for Language Science, Pennsylvania State University, University Park, PA 16802 United States
b Department of Modern Languages and Literatures, Fordham University, Bronx, NY 10458 United States

ARTICLE INFO

Article history:
Received 10 December 2015
Received in revised form 30 November 2016
Accepted 3 December 2016

Keywords:
Comprehension
Bilingualism
ERPs
Accent

ABSTRACT

This study tested semantic and grammatical processing of native- and foreign-accented speech. Monolinguals with little experience with foreign-accented speech listened to sentences spoken by foreign-accented and native-accented speakers while their brain activity was recorded using EEG/ERPs. We gathered behavioral measures of sentence comprehension, language attitudes, and accent perception. Behavioral results showed that listeners were highly accurate in comprehending both native- and foreign-accented sentences. ERP results showed that grammatical and semantic violations elicited different neural responses in native versus foreign accented speech. Native-accented speech elicited a frontal negativity (Nref) for grammatical violations and a robust N400 for semantic violations. However, in foreign-accented speech only semantic (not grammatical) violations elicited an ERP effect, a late negativity. Closer inspection of listeners who did and who did not correctly identify the foreign accent revealed that listeners who identified the foreign accent showed ERP responses for both grammatical and semantic errors: an N400-like effect to grammatical errors and a late negativity to semantic errors. In contrast, listeners who did not correctly identify the foreign accent showed no ERP responses to grammatical errors in the foreign-accented condition, but did show a late negativity to semantic errors. These findings provide novel insights into understanding the effects of listener experience and foreign-accented speaker identity on the neural correlates of language processing.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

With increased globalization, listeners and speakers from different language backgrounds are interacting with greater frequency, and the predominance of worldwide multilingualism (Marian & Shook, 2012) implies that such interactions often involve at least one non-native speaker of the language at hand. One of the most salient characteristics of a non-native speaker is a foreign accent (Gluszek & Dovidio, 2010). For a listener, foreign accents constitute a particularly challenging example of variability in the speech signal (Bent & Holt, 2013), often attributable to the speakers’ non-native pronunciation of words as well as speech rhythm and intonation patterns which deviate from the listener’s native language (cf. Reinsch & Weber, 2012). This variability induces intelligibility and comprehension difficulties (Munro & Derwing, 1995a, 1995b), but

* Corresponding author. Department of Modern Languages and Literatures, Fordham University, Bronx, NY 10458 United States.
E-mail addresses: sgrey4@fordham.edu (S. Grey), jgv3@psu.edu (J.G. van Hell).

http://dx.doi.org/10.1016/j.jneuroling.2016.12.001
0911-6044/© 2016 Elsevier Ltd. All rights reserved.
a respectable body of research demonstrates that listeners can quickly adapt to foreign-accented speech and that comprehension generally improves over time (for a review see Cristia et al., 2012).

Though behavioral research generally shows converging and comparable performance for foreign- and native-accented speech comprehension, recent neurocognitive studies using event-related potentials (ERPs) have produced divergent findings regarding the neural effects of foreign-accented language processing. An ERP study that tested native British English listeners’ semantic processing during sentence comprehension compared ERP effects to words spoken in a foreign, regional, or native accent (all correct items; Goslin, Duffy, & Floccia, 2012). The authors found that the N400 effect (an index of lexical/semantic processing) was reduced to words spoken in a foreign accent compared to a regional or native accent, which were not different from each other. Goslin et al. (2012) concluded that the unreliable phonetic variation inherent to foreign-accented speech (as opposed to reliable variation in native-accented speech) leads to reduced lexical activation and increased reliance on top-down contextual cues during comprehension.

Romero-Rivas, Martin, and Costa (2015) tested the effects of foreign-accented speech on semantic processing by examining semantic integration (N400 effects) and meaning reanalysis (P600 effects) in response to semantic anomalies (e.g., I drink my coffee with cream and *concrete before work each morning; *marks the anomaly). In their study, native Spanish listeners showed larger and more broadly-distributed N400s to semantic anomalies in foreign-compared to native-accented speech. This contrasts with Goslin et al. (2012) who tested all correct sentences as opposed to semantic anomalies and found reduced N400s. Additionally, Romero-Rivas et al. (2015) found that semantic anomalies elicited P600s in the native-accented but not foreign-accented condition. Romero-Rivas et al. interpreted the larger N400 as evidence that semantic anomalies were harder to process during foreign-accented speech comprehension, and suggest that this may be due to increased demands on lexical processing for the retrieval of the anomalous words. Regarding the absence of a P600 for foreign-accented semantic anomalies, the authors suggest that listeners avoid attempting to find an alternative meaning for the semantic violation when it is produced by a foreign-accented speaker, or that listeners do not have sufficient processing resources to carry out the reanalysis of meaning during foreign-accented speech comprehension.

In contrast to semantic anomaly processing, a related study by Romero-Rivas, Martin, and Costa (2016) tested the effects of foreign-accented speech on anticipatory lexical processing in correct semantics that varied by relatedness/expectancy. In the study, a group of native Spanish listeners heard foreign-accented sentences (Experiment 2) that contained a word corresponding to (1) the best lexical fit for the context (e.g., In the pirates’ map there was an X showing the location of the treasure; English translation, underlined in the original), (2) a semantically related, low cloze probability item (e.g., In the pirates’ map there was an X showing the location of the chest), or (3) a semantically unrelated, but plausible, low cloze probability item (e.g., In the pirates’ map there was an X showing the location of the enemy; examples from Romero-Rivas et al., 2016). The authors compared these data to a separate group of native Spanish listeners who were presented with the same sentences but spoken by a native-accented Spanish speaker (Experiment 1) and found equivalent N400 responses to sentences like (1), suggesting that native listeners show similar anticipatory lexical processes for native- and foreign-accented speech. They also observed larger N400s to sentences like (3) versus (2) in Experiment 1 (native-accented speech) but no such N400 amplitude difference in Experiment 2 (foreign-accented speech). The authors interpret this as evidence that although listeners can use equivalent anticipatory processes for native- and foreign-accented speakers during semantic processing, when the semantic expectations are not met (i.e., are not the best fit) the anticipatory processes do not show the same benefit for foreign-accented as compared to native-accented speech.

Hanulikova, Van Alphen, Van Goch, and Weber (2012) tested both semantic anomaly and grammatical error processing. In their study, which tested native Dutch listeners, N400s to semantic anomalies were not different in magnitude, but were more broadly distributed for Turkish-Dutch foreign-accented speech compared to native Dutch-accented speech. For grammatical processing, the authors examined P600 effects as an index of syntactic repair for Dutch agreement errors compared to correct agreement in sentences produced by the foreign- and native-accented Dutch speakers. Their findings showed that native-accented grammar errors elicited a P600 at the beginning of the experiment whereas foreign-accented errors did not elicit any comparable ERP effects. Hanulikova et al. (2012) interpreted the equivalently-sized N400s as evidence that the effect “was not modulated by the accent of the speaker” (p. 884) and that the listeners did not experience shallow processing or comprehension difficulties; the authors did not discuss implications of the broader N400 distribution for foreign-accented speech comprehension. For grammatical processing, Hanulikova et al. (2012) concluded that the listeners’ familiarity with the Turkish-Dutch foreign accent led them to modify their expectations about syntactic well-formedness for this foreign-accented speaker identity, since the error was common for those foreign-accented speakers. Indeed, when prompted to identify the foreign accent at the end of the experiment, over 80% of the listeners correctly identified the accent as Turkish.

In sum, all four studies used ERPs to test semantic processing of foreign-accented speech and each study found different patterns of results. Although Hanulikova et al. (2012) and Romero-Rivas et al. (2015) both found more broadly distributed N400s for semantic processing, Hanulikova et al. (2012) report comparably-sized N400 effects between foreign- and native-accented speech while Romero-Rivas et al. (2015) report larger N400s for foreign-than native-accented speech. Goslin et al. (2012) found smaller N400s for foreign compared to either native or regional accents, and the effects showed similar distributions. Finally, Romero-Rivas et al. (2016) found equivalent N400s for anticipatory semantic processing for expected lexical items in native- and foreign-accented speech, but their comparisons were made between-subjects. Moreover, both Goslin et al. (2012) and Romero-Rivas et al. (2016) tested all correct items, which differs from the semantic anomaly paradigm employed by the Hanulikova et al. (2012) and Romero-Rivas et al. (2015). Of the four studies, only Hanulikova et al. (2012)
tested grammatical processing in addition to semantics, and grammatical processing was different for native- and foreign-accented speech, with no neural effects being reported for foreign-accented grammatical errors.

One explanation for the divergent findings may be related to the types of listeners tested in the studies. Listeners vary in their own linguistic experience of being a non-native speaker of a foreign language (i.e., being bi-/multilingual or monolingual) and also in their extent of experience listening to foreign-accented speech. Listener experience with foreign-accented speech can affect the degree to which listeners adapt to foreign-accented input. Witteman, Weber, and McQueen (2013), for example, found that native Dutch listeners who had limited prior experience with German-accented Dutch showed lexical priming with medium and weakly accented words, but not strongly accented words. Listeners who had extensive experience with German-accented Dutch showed priming for all three accent types. The authors suggested that listeners with ample prior experience with foreign-accented variants have stronger links between the foreign speech variants and the canonical native forms. Listeners’ experience being bi-/multilingual or monolingual has also been found to affect psycholinguistic behavior towards foreign-accented speech. In a lexical priming study, Weber, Di Betta, and McQueen (2014) found that bilingual and monolingual listeners showed similar adaptation to foreign-accented speech variants, but that bilinguals were better than monolinguals at adapting to arbitrary accent features. The authors suggested that bilingualism enhances flexibility in phonetic-to-lexical mapping, which enables word recognition across different types of foreign-accented speech.

Returning to the four ERP experiments above, the listeners in Hanulikova et al. (2012) had extensive experience with the foreign accent tested in the study. Moreover, they were bi- or multilingual native listeners, and had extensive experience producing foreign-accented speech themselves. Romero-Rivas et al. (2015; 2016) also tested bilingual native listeners, though these listeners’ experience with foreign-accented speakers is unclear. Finally, Goslin et al. (2012) tested native English monolingual listeners who, given their monolingual status, likely had little experience producing foreign-accented speech themselves; their experience with foreign-accented speakers is unclear. Aside from all being native listeners, the participants in each of these studies represent different types of listeners, either by their experience with foreign-accented speakers, their monolingual/bilingual status, or both. Therefore, on the basis of the currently available evidence it is difficult to draw clear conclusions about the effects of foreign-accented speech on the neural basis of listeners’ real-time language processing.

In the present study, we isolated listener experience and monolingual/bilingual status by using ERPs to test semantic and grammatical processing of native- and foreign-accented speech in monolingual listeners who had limited experience with comprehension and production of foreign-accented speech.

In addition to differences in the types of listeners examined in these studies, other influential factors may have been at play. Effects of a foreign accent interact with social (Fuertes, Gottdiener, Martin, Gilbert, & Giles, 2012), cognitive (Lev-Ari & Keysar, 2012), and affective (Hatzidaki, Baus, & Costa, 2015) factors. The complexity and reach of the effects of foreign accent on language and communication are striking, but not well-understood. To begin to shed light on the influence of some of these factors, we examined the role of socially-relevant information on native- and foreign-accented speech comprehension.

Sociolinguistic and social psychology research demonstrates that foreign accentedness affects listeners’ attitudes, evaluations, and biases (e.g., Fuertes et al., 2012; Giles & Billings, 2004; Kinzler, Corriveau, & Harris, 2011). It is unknown whether these social underpinnings of foreign-accented speaker identity influence the neural correlates of foreign-accented speech comprehension. Intriguingly, related research shows that social-pragmatic information about a speaker—such as speaker sex (male/female), age (old/young), and socioeconomic status (high class/low class)—affects ERP responses for spoken language comprehension (Van Berkum, 2012), and ERP responses have been associated with individual differences in socially-relevant information, namely listener empathy (van den Brink et al., 2012). This research indicates that speaker identity is a pragmatic cue that affects language comprehension, and that individual differences in social factors affect comprehension outcomes. In the current study, we examined whether attitudes towards foreign accents influence the neurocognition of foreign-accented and native-accented language comprehension. To our knowledge, this is the first study to bridge sociolinguistic research on foreign accent and attitudes with research on the neural correlates on foreign-accented sentence comprehension (for recent discussion of social processing of spoken words, see Sumner, 2015; Sumner, Kim, & King, 2013).

To summarize, in the present study we used ERPs to investigate semantic and grammatical processing during foreign- and native-accented sentence comprehension in monolingual native listeners who had limited prior experience with foreign-accented speech. Additionally, we investigated the potential role of inter-individual variability in language attitudes in the processing of foreign- and native-accented speech.

2. Method

2.1. Participants

Participants were 39 monolingual native English listeners, all college students living in central Pennsylvania, USA, a highly monolingual context. Ten participants were excluded from analysis due to excessive artifact in the raw electroencephalogram (EEG) data (7), testing-session errors (2), or neuropsychological disorder (1). Thus, data from 29 monolingual listeners (mean age: 18.6 years, range 18–21; 5 male) are included in the analyses. This sample size fits with previous ERP studies on language processing and in particular ERP research on foreign-accented speech comprehension (Hanulikova et al., 2012; Romero-Rivas et al., 2015). None of the participants reported a history of neuropsychological, learning, or hearing disabilities. All were right-handed as assessed by an abridged version of the Edinburgh Handedness Inventory (Oldfield, 1971). Finally, none of the
participants were studying a foreign language at time of testing and none reported daily interactions with foreign-accented speakers. We therefore classify them as having limited experience with comprehension or production of foreign-accented speech (Wittman et al., 2013). These listeners were all fluent and proficient in English (verbal fluency $M = 43.66, SD = 3.7$; Michigan English Language Institute College Entrance Test, MELICET $M = 47.82, SD = 7.92$, max raw score = 50) and had normal IQs (Raven's Standard Progressive Matrices $M = 48.86, SD = 5.02$, max raw score = 60).

2.2. Materials

2.2.1. Stimuli and sentence listening task

The stimuli in the study consisted of declarative sentences that were grammatically and semantically well-formed, or had an error in English subject pronouns (he/she; example 1a below) or semantics (i.e., were semantically anomalous; example 1b below).

(1) a. Grammar target: Thomas was planning to attend the meeting but he/*she missed the bus to school.
   b. Semantic target: Kaitlyn traveled across the ocean in a plane/*cactus to attend the conference.

All sentences were pre-recorded by two female speakers, one with a native accent (standard American English) and one with a foreign accent (Chinese-English accent). The speakers were chosen based on accent ratings from a separate group of 23 native English listeners. The native English-accented speaker had a mean accent rating of 1.7 ($SD = 0.77$; 95% CI [1.4, 2.0]) on a Likert scale where 1 = no accent and 7 = very strong accent; the Chinese-English foreign-accented speaker had a mean rating of 5.6 ($SD = 0.92$; 95% CI [5.2, 5.9]). These ratings were significantly different, $t(44) = 15.45$, $p < 0.001$, $d = 4.3$.

Like Hanulikova et al. (2012), who examined the common Turkish-Dutch speaker production error of grammatical gender agreement, we focused on a common grammatical error for our target foreign-accented condition of Chinese-English accent speech: English subject pronouns. This grammar target is an attested production error in native Chinese speakers of English, even at high English proficiency and after years of living in English-speaking contexts (e.g., Johnson & Newport, 1989).

Each speaker recorded all correct, grammatical error, and semantic error sentences (480 sentences per speaker). Mean sentence duration for the native-accented speaker was 3234 ms ($SD = 486, 95\% CI [3190, 3277]$) and 4384 ms ($SD = 791, 95\% CI [4313, 4455]$) for the foreign-accented speaker; this difference was significant with a medium-to-large effect size, $F(1,479) = 765.75, p < 0.001$, $\eta^2_p = 0.62$. For the critical grammar items, mean native speaker pronoun duration in correct sentences was 117 ms ($SD = 33, 95\% CI [111, 123]$) and mean duration in pronoun error sentences was 121 ms ($SD = 35, 95\% CI [115, 123]$); mean foreign-accented speaker pronoun duration in correct sentences was 233 ms ($SD = 82, 95\% CI [218, 247]$) and mean pronoun duration in error sentences was 234 ms ($SD = 79, 95\% CI [219, 248]$). An ANOVA on durations with Accent (native, foreign) and Well-formedness (correct, error) as within-subjects factors showed that foreign-accented pronouns were longer in duration than native-accented pronouns (main effect of Accent, $F(1,119) = 706.62, p < 0.001$, $\eta^2_p = 0.85$) with no other significant duration effects (Well-formedness, $p = 0.698$, $\eta^2_p < 0.01$; Accent x Well-formedness, $p = 0.677$, $\eta^2_p < 0.01$).

For the critical semantic items, correct and anomaly words were matched on spoken word frequency in American English, $F(1,1239) = 2.03, p = 0.156$, $\eta^2_p = 0.01$. For the correct grammar items, mean native speaker pronoun duration in correct sentences was 117 ms ($SD = 39, 95\% CI [37, 41]$) and mean duration in pronoun error sentences was 148 ms ($SD = 35, 95\% CI [32, 38]$); mean foreign-accented speaker pronoun duration in correct sentences was 233 ms ($SD = 82, 95\% CI [218, 247]$) and mean pronoun duration in error sentences was 234 ms ($SD = 79, 95\% CI [219, 248]$). An ANOVA on durations with Accent (native, foreign) and Well-formedness (correct, error) as within-subjects factors showed that foreign-accented pronouns were longer in duration than native-accented pronouns (main effect of Accent, $F(1,119) = 706.62, p < 0.001$, $\eta^2_p = 0.85$) with no other significant duration effects (Well-formedness, $p = 0.698$, $\eta^2_p < 0.01$; Accent x Well-formedness, $p = 0.677$, $\eta^2_p < 0.01$).

For the critical semantic items, correct and anomaly words were matched on spoken word frequency in American English, $F(1,1239) = 2.03, p = 0.156$, $\eta^2_p = 0.01$. For the correct grammar items, mean native speaker pronoun duration in correct sentences was 117 ms ($SD = 39, 95\% CI [37, 41]$) and mean duration in pronoun error sentences was 148 ms ($SD = 35, 95\% CI [32, 38]$); mean foreign-accented speaker pronoun duration in correct sentences was 233 ms ($SD = 82, 95\% CI [218, 247]$) and mean pronoun duration in error sentences was 234 ms ($SD = 79, 95\% CI [219, 248]$). An ANOVA on durations with Accent (native, foreign) and Well-formedness (correct, error) as within-subjects factors showed that foreign-accented pronouns were longer in duration than native-accented pronouns (main effect of Accent, $F(1,119) = 706.62, p < 0.001$, $\eta^2_p = 0.85$) with no other significant duration effects (Well-formedness, $p = 0.698$, $\eta^2_p < 0.01$; Accent x Well-formedness, $p = 0.677$, $\eta^2_p < 0.01$).

For the critical semantic items, correct and anomaly words were matched on spoken word frequency in American English, $F(1,1239) = 2.03, p = 0.156$, $\eta^2_p = 0.01$. For the correct grammar items, mean native speaker pronoun duration in correct sentences was 117 ms ($SD = 39, 95\% CI [37, 41]$) and mean duration in pronoun error sentences was 148 ms ($SD = 35, 95\% CI [32, 38]$); mean foreign-accented speaker pronoun duration in correct sentences was 233 ms ($SD = 82, 95\% CI [218, 247]$) and mean pronoun duration in error sentences was 234 ms ($SD = 79, 95\% CI [219, 248]$). An ANOVA on durations with Accent (native, foreign) and Well-formedness (correct, error) as within-subjects factors showed that foreign-accented pronouns were longer in duration than native-accented pronouns (main effect of Accent, $F(1,119) = 706.62, p < 0.001$, $\eta^2_p = 0.85$) with no other significant duration effects (Well-formedness, $p = 0.698$, $\eta^2_p < 0.01$; Accent x Well-formedness, $p = 0.677$, $\eta^2_p < 0.01$).

For the critical semantic items, correct and anomaly words were matched on spoken word frequency in American English, $F(1,1239) = 2.03, p = 0.156$, $\eta^2_p = 0.01$. For the correct grammar items, mean native speaker pronoun duration in correct sentences was 117 ms ($SD = 39, 95\% CI [37, 41]$) and mean duration in pronoun error sentences was 148 ms ($SD = 35, 95\% CI [32, 38]$); mean foreign-accented speaker pronoun duration in correct sentences was 233 ms ($SD = 82, 95\% CI [218, 247]$) and mean pronoun duration in error sentences was 234 ms ($SD = 79, 95\% CI [219, 248]$). An ANOVA on durations with Accent (native, foreign) and Well-formedness (correct, error) as within-subjects factors showed that foreign-accented pronouns were longer in duration than native-accented pronouns (main effect of Accent, $F(1,119) = 706.62, p < 0.001$, $\eta^2_p = 0.85$) with no other significant duration effects (Well-formedness, $p = 0.698$, $\eta^2_p < 0.01$; Accent x Well-formedness, $p = 0.677$, $\eta^2_p < 0.01$).
Accuracy on the sub-types of stimulus items was also very high. Table 1 summarizes accuracy performance in more detail and includes the sub-types of stimulus items: correct and error, grammar and semantics for native-accented and foreign-accented conditions. For pronouns, an ANOVA with Accent (native, foreign) and Item (Correct, Error) showed no significant effect of Item \( p = 0.201, \eta^2_p = 0.10 \), indicating comparable accuracy in correctly labeling a correct pronoun sentence as correct and labeling a sentence containing a pronoun error as having a grammar error. The ANOVA also showed a significant effect of Accent: listeners were more accurate on native-accented than foreign-accented items, \( F(1,16) = 23.36, p < 0.001, \eta^2_p = 0.59 \). Finally, there was a significant Accent \( \times \) Item interaction, \( F(1,16) = 6.89, p = 0.018, \eta^2_p = 0.30 \), driven by better accuracy on correct foreign-accented than incorrect foreign-accented pronoun items (with native-accented correct and error items showing nearly matching accuracy). For semantics, the ANOVA showed no significant effect of Item \( p = 0.438, \eta^2_p = 0.04 \), indicating comparable accuracy in correctly labeling a correct semantic sentence as correct and labeling a sentence containing a semantic error as having a semantic error. Like pronouns and overall accuracy, there was a significant effect of Accent with a large effect size, \( F(1,16) = 70.72, p < 0.001, \eta^2_p = 0.82 \) whereby accuracy on native-accented items was better than accuracy on foreign-accented items. There were no other significant effects (Item \( \times \) Accuracy interaction, \( p = 0.296, \eta^2_p = 0.07 \)). To summarize, the error detection task administered to this separate group of 17 native English listeners showed that accuracy was better for native than foreign-accented items. This is to be expected based on research demonstrating that foreign-accented speech comprehension is challenging (e.g., Anderson-Hsieh & Koehler, 1988). Critically, the error detection task confirmed that grammar, semantic, and matched grammar/semantic correct items were detectable and perceived as grammar or semantic errors (or correct items) across both accent conditions.

During the EEG task, experimental sentences were presented in the context of a sentence listening task. Participants were told that they were going to listen to two people talk about their friends’ lives and were introduced to the names of 10 friends (5 male, 5 female names); no mention was made of foreign accent, grammar, or semantics. Each trial began with a 500 ms fixation cross in the center of a black screen followed by an auditorally-presented sentence delivered bi-aurally using earphones (Etymotic earphones, model ER4S; Elk Grove Village, IL), during which the fixation cross remained on the screen. After the sentence, participants either saw the word “Ready?” or saw a written comprehension question to which they responded using a button-box (response hand counter-balanced across lists and participants). The “Ready?” screen was designed to allow participants to blink or rest their eyes before initiating the next sentence. The comprehension questions served as a measure of comprehension performance and were additionally useful for keeping participants alert during the experiment. There were 48 yes/no comprehension questions for each list; 24 for each speaker identity condition; 12 required a yes response and 12 required a no response.

### 2.2.2. Language attitude survey

A language attitude survey was constructed to assess explicit attitudes towards foreign languages and accents. Such survey instruments are commonly used in sociolinguistic and applied linguistics research on attitudes (Dewaele & McCloskey, 2014; Garrett, 2010). The survey was piloted on a separate group of 20 English native speakers prior to being employed in the present study and consisted of 30 statements (Cronbach’s alpha = 0.89) such as “When I hear someone speak with an accent that is different from my own, I expect to have difficulty understanding them” and “I would work well in a professional or school environment with a person who speaks with a foreign accent.” This survey was administered to the 29 participants tested in the EEG study and they rated their strength of agreement with each statement where 1 = do not agree at all and 10 = agree completely.

### 2.2.3. Debriefing survey

Following the sentence listening task, participants were given a debriefing survey that asked whether they detected a difference in the accent of the two speakers. If participants selected ‘yes’, they were then asked to identify the accent of each speaker. This portion of the survey was ‘free response’ meaning that participants could type any response with respect to

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descriptive information for performance on the explicit error detection task (norming study).</td>
</tr>
<tr>
<td>N = 17</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td><strong>Native-accented speaker</strong></td>
</tr>
<tr>
<td>Grammar correct</td>
</tr>
<tr>
<td>Grammar error</td>
</tr>
<tr>
<td>Semantic correct</td>
</tr>
<tr>
<td>Semantic anomaly</td>
</tr>
<tr>
<td>Overall</td>
</tr>
<tr>
<td><strong>Foreign-accented speaker</strong></td>
</tr>
<tr>
<td>Grammar correct</td>
</tr>
<tr>
<td>Grammar error</td>
</tr>
<tr>
<td>Semantic correct</td>
</tr>
<tr>
<td>Semantic anomaly</td>
</tr>
<tr>
<td>Overall</td>
</tr>
</tbody>
</table>

Note. M = mean accuracy; SD = standard deviation; CI = confidence interval; maximum accuracy = 1.0.
accent identity of the speaker(s). If participants selected ‘no’ and therefore did not detect a difference in accent, they were prompted to submit the survey. Following accent identification for ‘yes’ responses, participants rated the degree of each accent (1–7 rating; 1 = no accent, 7 = very strong accent) and how easy it was to understand (1–7 rating; 1 = very easy to understand, 7 = very difficult to understand).

2.3. Procedure

Participants were tested in a single session lasting approximately 2.5 h and the study was approved by the university’s Institutional Review Board. After providing informed consent, participants completed a detailed background survey that assessed lifelong language experience, handedness, neuropsychological background, and socio-demographic information. After this, participants were seated in a comfortable chair in a sound-attenuated chamber. Participants were read aloud the instructions of the sentence listening task and completed practice prior to the experimental task. Participants were asked to relax and minimize eye-movements and blinks while listening to the sentences. As mentioned previously, participants were given an opportunity to blink or rest their eyes after each sentence (e.g., during the “Ready?” screen or written comprehension questions). Following EEG recording, participants completed a suite of cognitive and language tasks including the debriefing, a working memory task (automated operation span task, O-span; Unsworth, Heitz, Schrock, & Engle, 2005), an arrow-based Flanker task (Eriksen & Eriksen, 1974), verbal fluency (Luo, Luk, & Bialystok, 2010), the MELICET English proficiency test, and Raven’s Standard Progressive Matrices (Raven & Court, 1998).

2.4. ERP components

As described in the Introduction, Hanulikova et al. (2012) and Romero-Rivas et al. (2015) examined semantic processing by testing ERP responses to semantic violations (called anomalies) compared to correct items, and Hanulikova et al. (2012) additionally tested grammar processing by comparing grammatical gender agreement violations in Dutch to correct items. This violation paradigm — measuring the electrophysiological brain response to semantic or grammatical violations compared to correct items - is well-established in ERP language processing research (Kaan, 2007; Kutas & Federmeier, 2011) and reveals information on the neurocognitive correlates of language processing.

For semantic processing, such violations tend to elicit a centro-parietal negativity around 300–500 ms post-stimulus onset (i.e., after the onset of the correct or violation word). This is termed the N400 effect and is generally considered to reflect lexical/semantic access and integration processes (for a thorough review see Kutas & Federmeier, 2011).

For syntactic processing, the P600 is a common ERP effect found in response to violations of grammatical structure (e.g., Osterhout & Nicol, 1999; Osterhout, Kim, & Kuperberg, 2012), such as the Dutch gender agreement violations examined in Hanulikova et al. (2012). P600s are characterized by a predominantly posterior scalp distribution and tend to occur around 500–900 ms post-stimulus onset. The P600 effect is generally accepted as reflecting reanalysis or repair processes during language comprehension (e.g., Kaan, 2007; Osterhout et al., 2012). In the present study, we tested grammatical processing of pronouns, which have been found to elicit P600s in some studies (e.g., Filik, Sanford, & Leuthold, 2008; Osterhout & Mobley, 1995) but in other studies tend to elicit frontal negativities in the range of 270 ms through 1500 ms post-stimulus (Nieuwland & Van Berkum, 2006; Nieuwland, 2014; Streb, Rösler, & Hennighausen, 1999). This frontal negativity, called an Nref, is considered functionally distinct from P600 and N400 responses (Filik et al., 2008; Nieuwland & Van Berkum, 2006; Nieuwland, 2014) and believed to reflect attempts at resolving referential ambiguity or searching in memory for the proper antecedent (Hammer, Jansma, Lamers, & Munte, 2008; Nieuwland, 2014; Van Berkum, Brown, Hagoort, & Zwitserlood, 2003; Van Berkum, Zwitserlood, Bastiaansen, Brown, & Hagoort, 2004). The ERP components of interest in the present study, then, are the N400 to test semantic processing and the Nref (or P600) to test grammatical processing of pronouns.

2.5. EEG acquisition and analysis

During the sentence listening task, scalp EEG was recorded at a sampling rate of 500 Hz from 32 Ag/AgCl active electrodes (extended 10–20 system; Jasper, 1958) mounted in an elastic cap (Brain Products ActiCap, Germany). EEG was amplified using a Neuroscan Synamps RT system; it was filtered online with a 0.05–100 Hz bandpass filter and off-line with a 30 Hz half-amplitude low-pass filter (24 dB/octave roll-off). Scalp electrodes were referenced online to a vertex reference and re-referenced off-line to the average of activity recorded over the left and right mastoids. Additional electrodes were placed above and below the left eye and at the outer canthus of each eye, both referenced in bipolar montages, in order to screen for ocular artifacts. Impedances were kept below 5 kΩ.

ERPs, which were time-locked to the onset of the critical word for each sentence, were averaged off-line for both linguistic target conditions (grammar, semantics) in each participant (200 ms prestimulus baseline) within each speaker identity condition (native-accented, foreign-accented). Data free of ocular and muscle artifacts were included in the analyses. A total of 4.8% of trials were excluded due to artifacts. Informed by previous research, a time-window of 300–500 ms was selected to capture N400 and Nref effects (e.g., Nieuwland, 2014; Osterhout & Nicol, 1999) and a time-window of 500–900 ms was selected because of the sustained nature the Nref (e.g., Nieuwland, 2014; Van Berkum, Koornneef, Otten, & Nieuwland, 2007).
and to capture possible P600 effects (e.g., Bornkessel, Fiebach, & Friederici, 2004), which, as reviewed above, have also been found in response to pronoun mismatches. Data from the scalp electrodes were grouped into three distributional regions—anterior (Fz, F3, F4, FC5, FC6), central (Cz, C3, C4, CP1, CP2), and posterior (Pz, P3, P4, P7, P8) (for similar analytical approaches see Hanulikova et al., 2012; Romero-Rivas et al., 2015). Mean ERP amplitudes were entered into separate ANOVAs for each time window and linguistic target with Accent (native, foreign), Well-formedness (correct, error), and Distribution (anterior, central, and posterior) as within-subjects factors. We report Greenhouse-Geisser corrected $p$-values for data with more than one degree of freedom in the numerator. Follow-up analyses were conducted on significant interactions with the factor Accent or Well-formedness revealed in the global ANOVAs.

3. Results

3.1. Sentence comprehension

Over all questions, the listeners showed high comprehension accuracy, $M = 0.90, SD = 0.17, 95\% CI [0.83, 0.94]$. This was also true when the questions were divided by the two speaker conditions: for comprehension questions about sentences produced in the native-accented speaker condition, mean accuracy was $0.94, SD = 0.06, 95\% CI [0.92, 0.96]$ and mean accuracy on sentences in the foreign-accented condition was $0.91, SD = 0.06, 95\% CI [0.89, 0.94]$. Comprehension was high for both types of sentences, but it was slightly (3\%) higher for native-accented sentences than for foreign-accented sentences, $t(28) = 2.33, p = 0.027, d = 0.44$.

3.2. Debriefing

Recall that after the sentence listening task with EEG acquisition, we gave listeners a debriefing questionnaire that asked for information about their accent perceptions of the speakers they heard during the task. In this debriefing, 100\% of the listeners indicated they heard a difference in the accents of the speakers. Nearly 80\% of the listeners correctly identified the native-accented speaker as having an English accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian foreign accent$^1$ (Table 2). In contrast, only 36\% of listeners correctly reported an Asian

---

1 Although a correct identification rate of the native English accent of 80\% and an accent rating of 3.19 (out of 7) may seem odd on the surface, research indicates that native listeners vary in their categorizations/perceptions of even a native accent, and do not seem to show ceiling (100\%) correct identification (e.g., Goldstein, Knight, Bailis, & Conover, 1981; Preston, 1993).

2 Three people did not enter accent ratings, therefore the ‘ease of understanding’ and ‘degree of accent’ data reflect information from 26 listeners.
accent as Asian. The results from our behavioral measures also show that although sentence comprehension during the experimental task was high and therefore successful for both accent conditions (94% and 91% for native and foreign-accented speech, respectively), comprehension accuracy was slightly better for native-accented speech. Indeed, post-experiment subjective ratings showed the listeners reported experiencing a greater degree of comprehension difficulty for the foreign than for the native-accented speech.

3.3. ERP results

3.3.1. Grammar

Visual inspection of ERPs suggested that grammar errors in pronouns elicited a frontal negativity (Nref) for native-accented speech and no comparable ERP effects for the foreign-accented speech condition (Fig. 1). The results from the overall Accent (native, foreign) by Well-formedness (correct, error) by Distribution (anterior, central, and posterior) ANOVAs, for the 300–500 ms and 500–900 ms time windows, are presented in Table 3.

In the 300–500 ms time window, the ANOVA showed a significant main effect of Well-formedness (p = 0.018) and significant interactions for Accent x Distribution (p = 0.001) and Well-formedness x Distribution (p = 0.019). Follow-up analyses on the Well-formedness x Distribution interaction showed a significant effect of Well-formedness in the anterior (F(1,28) = 4.87, p = 0.005, \( \eta^2_p = 0.25 \)) and central (F(1,28) = 7.01, p = 0.013, \( \eta^2_p = 0.20 \)) regions (posterior, p = 0.68, \( \eta^2_p = 0.01 \)), indicating that pronoun errors elicited an Nref effect. This effect appeared to be driven by the ERP responses to the native-accented pronoun errors (Fig. 1). Follow-up analyses on the Accent x Distribution interaction to test the theory-driven a-priori hypothesis of differences in grammar processing for foreign-accented and native-accented speech further clarified this pattern. These analyses showed that only in the native-accented condition was there a significant effect at anterior/central sites: native-accented anterior, F(1,28) = 5.81, p = 0.023, \( \eta^2_p = 0.17 \); central, F(1,28) = 5.72, p = 0.024, \( \eta^2_p = 0.17 \) (posterior, F(1,28) = 0.04, p = 0.834, \( \eta^2_p = 0.002 \)); foreign-accented anterior, F(1,28) = 2.16, p = 0.152, \( \eta^2_p = 0.07 \); central, F(1,28) = 0.57, p = 0.455, \( \eta^2_p = 0.02 \); posterior, F(1,28) = 0.08, p = 0.775, \( \eta^2_p = 0.003 \).

In the 500–900 ms time-window (Table 3) there was a significant Well-formedness x Distribution interaction (p = 0.006) that was qualified by a 3-way Accent x Well-formedness x Distribution interaction (p = 0.025). Follow-up analysis on this
interaction showed a significant Well-formedness x Distribution interaction in the native-accented condition (F(2,56) = 12.78, p = 0.001, $\eta^2_g = 0.31$) that was driven by significant effects of Well-formedness in anterior, $F(1,28) = 4.38$, $p = 0.046$, $\eta^2_p = 0.14$, and also posterior regions, $F(1,28) = 7.046$, $p = 0.01$, $\eta^2_p = 0.20$ (central, $p = 0.55$, $\eta^2_p = 0.01$; in the foreign-accented condition the Well-formedness x Distribution interaction was not significant, $p = 0.295$, $\eta^2_p = 0.04$).

Thus, pronoun mismatches elicited a sustained N400 beginning in the 300–500 ms time-window and continuing to the 500–900 ms window when produced by the native-accented English speaker, and no comparable ERP effect when the same grammar error was produced by the foreign-accented English speaker.

3.3.2. Semantics

Visual inspection of the grand mean ERPs suggested that semantic errors elicited a large N400 effect for native-accented speech and a late negativity for foreign-accented speech$^3$ (Fig. 2). The results from the overall Accent (native, foreign) by Well-formedness (correct, error) by Distribution (anterior, central, and posterior) ANOVAs, for the 300–500 ms and 500–900 ms time-windows, are in Table 3.

In the 300–500 ms time window, the overall ANOVA showed a significant effect of Well-formedness ($p < 0.001$) and an Accent x Well-formedness interaction ($p = 0.004$). Follow-up analyses on this interaction showed that only in the native-accented condition did semantic errors elicit a significant effect of Well-formedness (N400 effect), $F(1,28) = 42.70$, $p < 0.001$, $\eta^2_p = 0.60$ (in the foreign-accented condition the effect of Well-formedness was not significant, $p = 0.106$, $\eta^2_p = 0.09$). Thus, semantic errors elicited an N400 response in the 300–500 ms time window for the native-accented speech condition and not the foreign-accented condition.

In the 500–900 ms time-window the overall ANOVA showed a significant effect of Well-formedness ($p < 0.001$) and a significant Well-formedness x Distribution interaction ($p = 0.012$). Follow-up analyses on this interaction showed that semantic errors elicited significant effects at all three distributional regions, and that this effect was more pronounced in the frontal and central regions than in the posterior region; anterior, $F(1,28) = 31.24$, $p < 0.001$, $\eta^2_p = 0.57$ (M $\mu V = -0.72$, $SE = 0.291$); central, $F(1,28) = 27.29$, $p < 0.001$, $\eta^2_p = 0.49$ (M $\mu V = -0.25$, $SE = 0.29$); posterior, $F(1,28) = 6.75$, $p = 0.015$, $\eta^2_p = 0.19$ (M $\mu V = 0.41$, $SE = 0.27$). So, in this later 500–900 ms time-window, semantic errors elicited a negativity across both accent conditions which appears to reflect the continuation of the large N400 for native-accented speech and the onset of late negativity for foreign accented speech with a predominantly anterior-to-central distribution (Fig. 2).

To summarize, ERP results indicate that grammatical and semantic errors elicited distinct neural responses when the errors were produced by a native-accented English speaker: an N400 for grammatical errors and an N400 effect for semantic errors. In contrast, only semantic errors, but not grammatical errors, elicited a neural response in the foreign-accented speaker condition, in the form of a late negativity.

3.3.3. Language attitudes and ERP effects

To investigate the relationship between listeners’ language attitudes and their neural responses to semantic and syntactic errors, we performed correlation analyses on the scores from the attitude survey and the ERP effect magnitudes. For the language attitude survey, the average score across all 30 items was entered as an overall language attitude score for each participant. The mean overall language attitude score was 6.49, SD = 1.02 95% CI [5.88, 6.59] (range 4.3–8.2; max. = 10,

$$^3$$ We also examined ERPs in the N400 time window to the word immediately following the critical semantic items for correct and anomaly sentences in native and foreign-accented conditions in order to test whether the tendency for semantic anomalies to be longer in duration had delayed consequences on processing. These analyses did not yield significant ERP effects, indicating that there were not delayed consequences on processing.
Of the 30 items, 12 were exclusively related to foreign accents, and the average score for those 12 items was entered as an accent attitude score. The mean accent attitude score was 6.86, SD = 1.23 95% CI [6.44, 7.30] (range 4.4–7.8).

Regarding the ERP effect magnitudes, we calculated each listener’s Nref effect magnitude (correct minus error) in the 300–500 ms time window over the anterior region (electrodes Fz, F3, F4, FC5, FC6). This accounts for the ERP response to native-accented grammatical errors. We also calculated each listener’s N400 effect magnitude for the semantic errors over a centro-parietal region (electrodes Cz, C3, C4, CP1, CP2, Pz, P3, P4, P7, P8) in the 300–500 ms and 500–900 ms time windows. This accounts for the N400 effect (native-accented condition) and later negativity (foreign-accented condition) observed in response to semantic anomalies.

The results from our correlation analyses are reported in Table 4. There were no significant correlations. These findings suggest that self-reports of language attitudes were not related to the magnitude of neural responses to grammatical or semantic errors in either speaker condition. We also tested whether any of the cognitive or language tasks we administered were related to ERP effect magnitudes, and no significant relationships were found (Table 4).

Table 4

<table>
<thead>
<tr>
<th></th>
<th>Native accent</th>
<th>Foreign accent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nref effect magnitude</td>
<td>N400 effect magnitude</td>
</tr>
<tr>
<td>Overall language attitude</td>
<td>0.176</td>
<td>-0.134</td>
</tr>
<tr>
<td>Accent attitude</td>
<td>0.273</td>
<td>-0.165</td>
</tr>
<tr>
<td>Flanker effect</td>
<td>0.270</td>
<td>-0.165</td>
</tr>
<tr>
<td>O-span</td>
<td>0.102</td>
<td>-0.016</td>
</tr>
<tr>
<td>Raven’s</td>
<td>-0.224</td>
<td>0.146</td>
</tr>
<tr>
<td>Verbal fluency</td>
<td>-0.153</td>
<td>0.284</td>
</tr>
<tr>
<td>MELICET</td>
<td>0.172</td>
<td>0.217</td>
</tr>
</tbody>
</table>

Note. All p’s > 0.10. Value reported is Pearson correlation coefficient. Raven’s = Raven’s Standard Progressive Matrices; MELICET = English proficiency test; O-span = working memory span test.
3.4. Post-hoc analysis: foreign-accented identity and ERP effects

As described above, our debriefing results showed that about a third (36%) of the listeners correctly identified the foreign-accented speaker as having an Asian accent. This is smaller than the 54% of monolingual listeners in Goslin et al. (2012) who correctly identified the target foreign accents and much smaller than the 80% of bilingual listeners in Hanulikova et al. (2012) who correctly identified the foreign accent. To investigate the effects of foreign accent identification in more detail, we conducted a follow-up ERP analysis on the listeners who correctly identified an Asian accent (n = 10) and those who identified the foreign-accented speaker as having an accent other than Asian (n = 18; Table 2). We conducted ANOVAs on mean ERP amplitudes for each target structure (grammar and semantics) in both of these sub-groups, with Well-formedness (correct, error) and Distribution (anterior, central, posterior) as within-subjects factors.

The ERP waveforms for grammatical and semantic processing of foreign-accented sentences in each sub-group are presented in Figs. 3 and 4. Visual inspection of the ERP waveforms indicated that grammatical errors elicited a broad negativity in the sub-group of listeners who correctly identified the Asian foreign accent (Fig. 3a), whereas there were no visually apparent ERP effects in the grammar condition for listeners who did not identify the Asian foreign accent (Fig. 4a). In contrast, the ERP responses to semantic errors appeared similar for both sub-groups of listeners: a late negativity (Figs. 3b and 4b).

Results from the Well-formedness by Distribution ANOVAs in 300–500 and 500–900 ms time windows for the sub-group of listeners who successfully identified the Asian foreign accent showed the following. For grammar processing, there was a significant main effect of Well-formedness in the 300–500 ms time window ($F(1,9) = 8.0, p = 0.02, \eta^2_p = 0.47$) and no significant ERP effects for grammar errors in the 500–900 ms window (Well-formedness, $p = 0.71, \eta^2_p = 0.02$; Well-formedness x Distribution, $p = 0.18, \eta^2_p = 0.18$). This result indicates that the listeners who correctly identified the Asian accent showed neural sensitivity to pronoun errors produced by the foreign-accented speaker. For semantic processing, there were no significant ERP effects in the 300–500 ms window (Well-formedness, $p = 0.28, \eta^2_p = 0.13$; Well-formedness x Distribution, $p = 0.76, \eta^2_p = 0.01$) but there was a significant effect of Well-formedness in the 500–900 ms time window, $F(1,9) = 9.22, p = 0.014, \eta^2_p = 0.51$. This result mirrors the larger group pattern for a late negativity in response to semantic errors produced by the foreign-accented speaker.

**Fig. 3.** Grammar and semantic processing of foreign-accented speech in the sub-group of listeners who correctly identified the Asian accent. Panel A: Grand mean ERP waveforms for the foreign-accented pronoun condition over midline electrodes for correct pronoun (black line) and pronoun error (blue solid line) conditions. Each tick mark represents 100 ms; negative voltage is plotted up. Topographical maps show the scalp distribution of activity in the grammar error minus correct conditions, averaged for the 300–500 and 300–900 ms time windows. Calibration scale is ±2 μV. Panel B: Grand mean ERP waveforms for the foreign-accented semantic condition over midline electrodes for correct semantics (black line) and semantic error (blue dashed line) conditions. Topographical maps show the scalp distribution of activity in the error minus correct conditions, averaged for the 300–500 and 500–900 ms time windows. Calibration scale is ±2 μV. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
ANOVA results from the sub-group of listeners who did not identify the Asian foreign accent showed no significant effects for grammar processing\(^4\) in either time window; 300–500 ms: Well-formedness, \(p = 0.85, \eta_p^2 < 0.01\); Well-formedness x Distribution, \(p = 0.17, \eta_p^2 = 0.11\); 500–900 ms: Well-formedness, \(p = 0.29, \eta_p^2 = 0.07\); Well-formedness x Distribution, \(p = 0.42, \eta_p^2 = 0.04\). For semantic processing, there were no significant ERP effects in the 300–500 ms time window (Well-formedness, \(p = 0.19, \eta_p^2 = 0.09\); Well-formedness x Distribution, \(p = 0.38, \eta_p^2 = 0.05\)) but there was a significant main effect of Well-formedness in the 500–900 window, \(F(1,17) = 9.86, p = 0.006, \eta_p^2 = 0.37\), which aligns with the larger group pattern for a late negativity in response to semantic errors in the foreign-accented condition.

In sum, closer inspection of listeners who did and who did not correctly identify the foreign accent revealed that listeners who identified the foreign accent showed ERP responses for both grammatical and semantic errors: an N400-like response for grammar and a late negativity to semantic errors. In contrast, listeners who did not correctly identify the foreign accent showed no ERP responses to grammatical errors in the foreign-accented condition, but did show the late negativity to semantic errors.

4. Discussion

In this study we used ERPs to examine grammatical and semantic processing of native-accented and foreign-accented sentences in monolingual native listeners. Behavioral results showed that listeners were highly accurate in both native- and foreign-accented sentence comprehension, although they were significantly more accurate in comprehension of native-

---

\(^4\) Instead of an absence of ERP effects here, one possibility is that the listeners showed an effect of pronoun mismatch on words following the pronoun. We examined the ERPs to the word immediately following the correct/error pronoun targets for the foreign-accented speech condition in this subset of listeners and did not find any significant ERP effects on this post-pronoun word.
than foreign-accented sentences. The ERP findings revealed further differences in real-time sentence processing. Grammatical and semantic errors elicited different neural responses for the native-accented and foreign-accented speaker conditions. The native-accented English speech elicited an Nref effect for grammatical errors in pronouns and a large N400 effect for semantic errors. However, in the foreign-accented speaker condition, only semantic, but not grammatical, errors elicited an ERP response, in the form of a late negativity.

We evaluated foreign-accented speech comprehension in more detail by examining the effect of correct identification of the foreign accent on the neural correlates of grammatical and semantic processing. Listeners who correctly identified the Asian foreign accent showed neural responses to both grammatical and semantic errors. Specifically, these listeners showed an N400-like response in the 300–500 ms range following grammatical errors produced by the foreign-accented speaker and showed a late negativity in the 500–900 ms range following semantic errors. In contrast, the listeners who identified the foreign accent as something other than Asian showed no ERP responses to grammatical errors produced in the foreign-accented condition, but did show the late negativity. Notably, these differences in neural processing between the two subgroups were not reflected in behavioral measures of global sentence comprehension, as both groups demonstrated comparably high comprehension accuracy of the foreign-accented sentences. A hypothesis put forth by Hahne and Friederici (2001) for late negativities versus N400s in a study on native and monolinguals listening to foreign-accented semantic anomalies provides new insight on the effects of foreign-accented speech on language processing. One interpretation of the late negativity is that it reflects a search in memory for the proper antecedent.
when the antecedent is missing (e.g., Nieuwland, 2014; Van Berkum et al., 2003). Thus, in the native-accented condition, listeners showed ERP evidence of antecedent search when encountering the mismatching pronoun.

A novel finding from the current study is that, for foreign-accented speech, the subset of listeners who correctly identified the foreign accent of the speaker also showed a negativity to foreign-accented pronoun mismatches, but this negativity was more consistent with an N400 effect. As reviewed in Nieuwland (2014), N400s in response to pronoun mismatches suggest difficulty in retrieving semantic information for the pronoun. The N400 effect found here for foreign-accented grammar errors appears to be smaller and more right-lateralized than the standard centro-parietal distribution for N400s. If, as we suggest above, semantic processing is more difficult for inexperienced monolingual listeners in foreign-accented speech contexts, then initiating retrieval of semantic information for the foreign-accented pronoun would also be more difficult. This could help explain the size and distribution of the N400 effect to foreign-accented pronoun mismatches in this subset of listeners. Critically, the finding of an Nref for native-accented pronouns but an N400 for foreign-accented pronouns indicates that the listeners employed functionally distinct processes for the same target grammatical structure as a function of native-versus foreign-accented speaker identity, at least the listeners who could identify the foreign accent.

The listeners who did not correctly identify the foreign accent, on the other hand, showed no ERP sensitivity to foreign-accented pronoun mismatches. Although prudence is needed when interpreting an absence of ERP effects, previous work that has observed ‘absent’ effects for foreign-accented speech can help interpret the result observed in this subgroup of listeners. Hanulikova et al. (2012) suggested that the absence of a P600 effect to foreign-accented grammatical agreement errors indicated that the listeners, who were highly familiar with the Turkish-Dutch foreign accent of the study, had modified their expectations about grammatical well-formedness for the foreign-accented speaker identity, since the error was common for that identity. Since the listeners in the current study were not highly familiar with the foreign accent, and a large subgroup could not even correctly identify it, the absence of ERP effects to foreign-accented pronoun mismatches in this subgroup is unlikely to be due to these listeners having modified their expectations about grammatical well-formedness for the speaker. Romero-Rivas et al. (2015) suggested that the absence of P600 effects for foreign-accented semantic anomalies indicated that the listeners avoided attempting to find an alternative meaning for the anomaly, or alternatively that they did not have enough processing resources to reanalyze the meaning in the foreign-accented speech context. Applying this account to the (absent) outcome for foreign-accented pronoun processing in the listeners who did not identify the foreign accent implies that they either avoided initiating antecedent search (Nref) for foreign-accented pronoun mismatches or avoided attempting retrieval of semantic information (N400) for the pronoun. Overall, the results from our subsets of listeners who could and who could not classify the identity of the foreign-accented speaker reveal that native-accented and foreign-accented speech, together with listeners’ ability to identify the foreign accent of a speaker, have consequences for the neural correlates of grammatical processing during speech comprehension.

The distinct neural patterns we observed for semantic and grammatical processing of native- and foreign-accented speech also help to understand the behavioral outcomes for sentence comprehension. Accuracy in sentence comprehension was very significantly better for native-accented sentences, whereas they did for native-accented mismatches, whereas they did for native-accented mismatches. For semantics, the listeners exhibited difficulty in lexical/semantic processing of foreign-accented speech (late negativity) while they showed intact and robust semantic processing of native-accented speech (classic N400 effect).

Though foreign-accented speech affects behavioral sentence comprehension and ERP outcomes, we did not find any relationship between social underpinnings of foreign-accented speaker identity, i.e., listeners’ language attitudes and attitudes towards foreign accents, and the observed neural outcomes. Foreign-accentedness has been found to affect listeners’ attitudes, personal and professional evaluations, and biases (e.g., Fuertes et al., 2012). The lack of significant relationships between ERPs to foreign-accented speech and language attitudes may be because the language attitude survey was a self-report measure. Though common in research on attitudes, self-report surveys reflect explicit attitudes and self-reporting can distort more subtle, or implicit, attitudes and biases (Pantos & Perkins, 2013). Possibly, a less explicit measure of attitudes (such as the Implicit Association Test, IAT; Greenwald, McGhee, & Schwartz, 1998) will reveal relationships between the neural correlates of foreign-accented speech comprehension and listeners’ attitudes towards foreign-accented speech.

As globalization expands and bilingualism becomes more prevalent, interactions among native and non-native language speakers who are likely to have a foreign accent will become increasingly common. In this study, we demonstrated that the neural underpinnings of real-time sentence comprehension are influenced by foreign-accented speaker identity, that the effects vary for semantic versus grammatical processing, and for monolingual listeners who can classify the identity of the foreign-accented speaker versus those who cannot classify this identity.

---

6 An N400 effect in this foreign-accented pronoun context could also be explained as an effect for unexpected lexical items, e.g., expecting ‘he’ but hearing ‘she’ instead. We thank one of the reviewers for pointing this out.
Funding acknowledgment

This research was supported by NSF OISE-0968369 and NSF BCS-1349110 to Janet van Hell and NSF SMA-1514276 to Sarah Grey and Janet van Hell.

Acknowledgments

We thank Courtney Johnson-Fowler, Yiran Zhang, Carrie Jackson, Kaitlyn Litofsky, Tim Poepsel, Kathleen Ammerman, Erika Exton, Christine Hughes, and Silviana Lee for assistance with this project. We thank Darren Tanner for especially valuable input on a previous version of the manuscript.

References


