Neural responses to morphosyntactic violations in foreign-accented speech

Rebecca Jane Holt

BA(Adv.)(Hons)

Macquarie University

Department of Linguistics

Thesis submitted in fulfilment of the requirements of the degree of

Master of Research (Linguistics)

10th October, 2016

Table of Contents

Abstract

When comprehending speech, listeners extract information regarding the speaker's characteristics, such as their sex, nationality or socioeconomic status, from the speech signal. This speaker-identity information is integrated on-line during speech processing, influencing how listeners process both the phonological and semantic content of the utterance. However, little is known as to whether morphosyntactic processing is also influenced by the integration of the speaker's identity. This study used the event-related potential technique to examine whether listeners' (n = 30) neural responses to morphosyntactic violations (here subject-verb agreement errors) were modulated by the presence of speaker-identity information in the form of a foreign accent.

The results demonstrated that the presence of a foreign accent did affect the neural responses observed for the violations, with errors of the commission type (where a superfluous morpheme was present) eliciting an N400 effect in native speech but an anterior negativity in foreign-accented speech. However, omission errors (where a required morpheme was missing) elicited no violation effect for either speaker. The unexpected presence of an N400 effect was attributed to shallow syntactic processing and a possible priming effect caused by the presence of semantic violations in the experimental stimuli. The lack of effect observed for omission errors was attributed to the relative lack of perceptual salience of this error type. Some individual variation in the neural responses was observed, which was weakly related to participants' empathy levels. Overall, the results provided evidence that speaker-identity information is indeed integrated on-line during morphosyntactic processing.

Declaration

I, Rebecca Jane Holt, do hereby declare that the research contained in this thesis has not been submitted for any other higher degree or to any other university or institution. I declare that I have made every effort to acknowledge the work and ideas of others and indicate the sources of information used in the text. This research study was approved by the Macquarie University Human Ethics Research Committee (reference number 5201600062).

Acknowledgements

I would like to express my deepest gratitude to a variety of people who have helped me in so many ways while writing this thesis. First and foremost, many thanks are due to my supervisors: Professor Katherine Demuth, Carmen Kung, and Dr Elaine Schmidt. I would particularly like to thank Katherine for her overall guidance of the project, Carmen for her patience and help in data collection and analysis, and Elaine for her expertise in stimuli design and preparation. I would also like to thank all the members of Macquarie University's Child Language Lab and Phonetics Lab, who have provided invaluable feedback and encouragement along the way. Thanks are also due to all the members of the Writing Group, for persevering through lengthy drafts and still providing such helpful feedback.

On a more personal note, I have to thank my parents for being so supportive of my academic endeavours in such a variety of ways (including, but not limited to, washing the large numbers of towels I would bring home after a run of consecutive EEG sessions – thanks Mum!), and my brother, Matthew, for putting up with having me in the house. Huge thanks are also due to my fiancé, Ben, for his encouragement, and for understanding that I didn't care about him any less in those times when I just needed to ignore him and work on my thesis. Thanks are due to my MRes colleagues, particularly Sujal Pokharel and Tünde Szalay, for encouraging and commiserating in equal measure. Finally I'd like to thank my church family for their support and prayers, and of course God, who has had everything under control all along.

List of tables

List of Figures

1. Introduction

The brain rapidly integrates information from a variety of sources during speech processing. This includes not only the content of the utterance itself, but information about the speaker who produced the utterance, such as the speaker's sex, socioeconomic status or ethnicity (hereafter 'speaker-identity information'). There is considerable evidence that the integration of this speaker-identity information influences how phonological and semantic information are processed (e.g. Johnson, Strand, & D'Imperio, 1999; Van Berkum, van den Brink, Tesink, Kos, & Hagoort, 2008). However, little is known as to whether and how speaker-identity information may affect syntactic or morphosyntactic processing.

So far, only one study has addressed the influence of speaker-identity information on morphosyntactic processing. Hanulikovà, van Alphen, van Goch and Weber (2012) examined neural responses to grammatical gender agreement violations in both native and Turkish-accented Dutch. This study found that speaker identity, as conveyed by the speakers' accents, did indeed modulate the processing of the morphosyntactic violations. The authors hypothesised that the observed effect was due to the listeners' expectation of morphosyntactic violations in Turkish-accented speech, built upon their previous experience with Turkish-accented speakers of Dutch.

These results raise the question of what differences may be observed between neural responses to errors which, according to the speaker-identity information presented, are highly expected and errors which are unexpected given the speaker's identity. The current study aimed to build upon the findings of Hanulikovà and colleagues by examining the integration of speaker-identity information during morphosyntactic processing in English. Subject-verb agreement errors in Mandarin-accented English were examined to determine whether or not the findings of Hanulikovà and colleagues' study generalise across languages and morphosyntactic violations. Additionally, the unique properties of subject-verb agreement errors in Mandarin-accented English enabled the investigation of the effect of violation expectedness on the listeners' neural responses. This allowed the current study to further examine the claim that observed differences in neural responses between native and foreign-accented speech are due to listeners' expectations based on prior experience.

This study utilised electroencephalography (EEG) and the event-related potential (ERP) technique to examine listeners' neural activity. Thus, to facilitate discussion of the relevant literature, this thesis will begin with a methodological review covering EEG and ERPs, with an emphasis on the language-related ERP components relevant to the study (sections 2.1 and 2.2). This will be followed by a review of the relevant theoretical issues (sections 2.3 to 2.5), culminating in the hypotheses of the current study (section 2.6). Section 3 will describe the methodology employed in the study, followed by the results (section 4), discussion (section 5) and conclusion (section 6).

2. Literature review

2.1. Electroencephalography and event-related potentials

One way in which researchers may gain an understanding of the cognitive and neural processes involved in language processing is by investigating patterns of brain activity which co-occur with linguistic stimuli. These neural correlates of language processing can provide insights into the occurrence and relative timing of various neuro-cognitive processes. They may also elucidate whether different types of linguistic information are processed in the same or different ways.

Electroencephalography (EEG) is a widely used method of investigating neural correlates by measuring the electrical activity occurring in the brain. Electrodes record voltage changes at the scalp resulting from the synchronised activity of large groups of neurons (see Luck, 2005). However, raw voltages drawn from EEG data are difficult to interpret and unlikely to be informative as they are influenced by a wide variety of random factors, such as individual differences in the skull and scalp of the participants, variability in the orientation of the dipole generating the electrical activity, or even whether or not the participant washed their hair on the day of the EEG (Cohen, 2014).

To address this issue, event-related potentials (ERPs) can be calculated from the EEG signal. ERPs are patterns of scalp-recorded neural electrical activity time-locked to particular events, such as the presentation of a stimulus or the execution of a motor response (Luck, 2005; Coulson, King, & Kutas, 1998). These are generally calculated by averaging together many trials, then subtracting the averaged waveforms between conditions. This is intended to cancel out the effect of the various random factors mentioned above, as well as any electrical activity that remains constant between conditions (i.e. is not generated in response to the stimulus of interest). The resulting ERPs can be characterised by their polarity, latency, amplitude and scalp distribution and can be correlated to specific cognitive behaviours associated with the processing of the stimulus.

EEG and the ERP technique have several benefits over other methodologies, such as functional magnetic resonance imaging (fMRI) and magnetoencephalography (MEG). The excellent temporal resolution of ERPs provides a distinct advantage over fMRI as a tool to investigate language processing, as it can provide insight into the exact time-course of cognitive processes. This allows ERPs to be linked to highly specific language events, such as particular semantic or syntactic violations, which would not be possible with fMRI data, as it has poor temporal resolution (Luck, 2005). MEG, like EEG, has excellent temporal resolution, and slightly better spatial resolution than EEG, however EEG is both more comfortable for participants and less expensive than MEG (Luck, 2005). Additionally, the majority of the existing body of literature regarding the neural correlates of speech processing is based on EEG studies. Thus by using EEG, the results of the current study will be more comparable to those of existing studies.

Despite these clear advantages, the ERP technique does suffer from some limitations. The spatial resolution of EEG is very poor due to the smearing of electrical signals as they are conducted through the brain, skull and scalp (Kappenman & Luck, 2012). Additionally, any observed topographical distribution of electrical activity at the scalp, assuming that it has not been smeared, could still theoretically be attributed to an infinite number of potential source configurations (Luck, 2005). For this reason, no claims can be made about the location in the brain of the neural generators of the EEG signal. Additionally, the interpretation of EEG data can be problematic as it is often difficult to link an observed EEG component to the exact cognitive function that it represents. While the ERP literature has provided evidence that certain components are linked to certain aspects of processing, the specific function that the component is correlated with often remains controversial (see section 2.2). This disadvantage can be mitigated, however, by using ERPs as a method of differentiating between neural responses to different stimuli. The current study will take this approach, examining ERP responses across conditions without making any claims about the precise nature of which cognitive processes the different responses represent.

2.2. Sentence-processing ERP components

There are three ERP components relevant to the current study: the N400, P600 and left anterior negativity (LAN). Though these components are not necessarily language-specific, emphasis in the following sections will be placed on their relevance to sentence-processing research.

2.2.1. N400

The N400 component (Kutas & Hillyard, 1980) is typically associated with semantic processing. It occurs as a negative shift in the ERP waveform approximately 400 ms after the onset of the stimulus and greatest over the centro-parietal scalp region (Swaab, Ledoux, Camblin, & Boudewyn, 2012). Latency of the N400 is generally stable (Kutas & Federmeier, 2011), although N400 responses to auditory stimuli may occur earlier than those in response to written stimuli, sometimes as early as 100 ms post-stimulus (Kutas, Neville, & Holcomb, 1987; Swaab et al., 2012). The topographical distribution of the N400 also shows a tendency to vary according to stimulus modality, with auditory N400s showing a more left-lateralised distribution than those observed in reading (e.g. Hagoort & Brown, 2000a).

The N400 has been observed in response to semantic information across a range of modalities, including spoken, signed and written language (Kutas et al., 1987). Typically, N400s are observed in response to all lexical items in an utterance. The amplitude of the N400 associated with a particular lexical item is inversely related to that item's context-based cloze probability (Kutas & Hillyard, 1984). That is, the largest N400s are elicited by lexical items which create an irresolvable semantic incongruity in context (extremely low cloze probability, e.g. "he took a sip from the *transmitter*"), while moderate N400s are elicited by lexical items which, though semantically plausible, are unexpected (moderately low cloze probability, e.g. "he took a sip from the *waterfall*") (Kutas & Federmeier, 2000). Additionally, some studies have reported N400-like effects occurring at the end of sentences which contain *syntactic* violations (e.g. Hagoort, 2003; Osterhout & Holcomb, 1992), perhaps reflecting difficulty in semantic integration of the sentence as a whole.

Recent accounts suggest that, rather than interpreting the N400 as a marker of semantic processing in language specifically, the N400 could better be described as a marker of meaning processing in general (Kutas & Federmeier, 2011). This is based on the fact that the N400 has also been observed in non-linguistic contexts, including the processing of drawings (Ganis, Kutas, & Sereno, 1996; Holcomb & McPherson, 1994), environmental sounds (van Petten & Rheinfelder, 1995) and mathematics (Galfano, Mazza, Angrilli, & Umiltà, 2004). Regardless of modality, the predominant functional accounts of the N400 link the component to either semantic memory access or semantic integration processes (see, e.g., Kutas & Federmeier, 2011).

2.2.2. P600

The P600 component (Osterhout & Holcomb, 1992) or Syntactic Positive Shift (Hagoort, Brown, & Groothusen, 1993) is typically associated with syntactic processing, although its functional significance is somewhat more controversial than that of the N400. This component is typically observed as a long, late positive shift in the waveform, generally peaking around 600 ms post-stimulus-onset (Swaab et al., 2012). It is often broadly distributed, though generally strongest posteriorly (Swaab et al., 2012). Variability in P600 latency and distribution has been observed, prompting some to propose a distinction between early and late P600s (Hagoort & Brown, 2000b; Molinaro, Barber, & Carreiras, 2011). According to this view, early P600s occur around 500-750 ms post-stimulus and are broadly distributed, even occurring at frontal scalp areas (and may indeed be limited to frontal scalp areas), while late P600s occur around 750-1000 ms post-stimulus and are restricted to the posterior regions of the scalp.

The P600 has been observed in response to a range of syntactic phenomena including tense, agreement, case, phrase structure and subcategorisation violations, garden-path sentences and sentences with wh-dependencies (Hagoort et al., 1993; Coulson et al., 1998; Friederici, Pfeifer, & Hahne, 1993; Gouvea, Phillips, Kazanina, & Poeppel, 2010). While the P600 is generally described as a violation response, irresolvable syntactic violations are not necessary for P600 occurrence: the P600 also occurs in response to congruous but difficult-to-process syntactic structures (Kaan, Harris, Gibson, & Holcomb, 2000; Gouvea et al., 2010). In addition, the P600 has been observed in response to sentences containing certain types of *semantic* violations, particularly thematic role violations (Frenzel, Schlesewsky, & Bornkessel-Schlesewsky, 2011).

Several functional interpretations of the P600 have been proposed, which can be grouped into three broad categories. Firstly, some interpretations treat the P600 as a response to syntactic violations or non-preferred syntactic structures. Generally, interpretations of this kind suggest that the P600 reflects a syntactic reanalysis or repair process (Friederici, 1995, 2002). Some suggest that the P600 may index a more general conflict-monitoring process that occurs in response to a mismatch between the outputs of different processing levels caused by a syntactic violation (Kuperberg, 2007), or it may reflect a categorisation process, classifying sentences as being syntactically ill-formed (Frenzel et al., 2011).

A second group of interpretations suggest that the P600 instead reflects the process of building syntactic structures. According to this view, the P600 may reflect the retrieval and integration of the elements of syntactic structures (Gouvea et al., 2010), or it may function as an index of syntactic integration difficulty, much as the N400 may index semantic integration difficulty (Kaan et al., 2000). The observation of the P600 in non-linguistic contexts, including music (Patel, Gibson, Ratner, Besson, & Holcomb, 1998) and mathematics (Nùñez-Peña & Honrubia-Serrano, 2004), has also prompted the suggestion that the P600 may in fact be a non-domain-specific indicator of 'structure building', rather than a marker of syntactic structure building specifically (Swaab et al., 2012). Another alternative view is that, as the P600 has also been observed for pragmatic violations, it may reflect increased effort in constructing a coherent representation of what is being communicated, rather than effort in constructing syntactic structures in particular (Hoeks, Stowe, Hendriks, & Brouwer, 2013).

Finally, others suggest that the P600 is a member of the P300 family of components, which typically occur in response to surprising (i.e. rare and/or salient), task-relevant stimuli (Donchin, 1981). According to this view, the P600 is related to the probability of certain syntactic constructions rather than syntactic processing *per se* (Coulson et al., 1998; Bornkessel-Schlesewsky et al., 2011).

2.2.3. Left anterior negativity

The left anterior negativity (LAN; Friederici et al., 1993), like the P600, is associated with elements of syntactic processing, though particularly morphosyntactic processing. This component occurs as a negative shift in the waveform, generally observed over the left anterior scalp region between 300 and 500 ms post-stimulus-onset. The LAN can also be bilaterally distributed in some cases, where it may be referred to as an anterior negativity (AN). The LAN is typically distinguished from the N400, which occurs at a similar latency, by its topographical distribution (left anterior versus centro-parietal; Molinaro et al., 2011), despite the fact that the distribution of the LAN is variable.

LANs have been observed in response to syntactic and morphosyntactic violations of various kinds, though predominantly agreement violations (e.g. Friederici et al., 1993; Molinaro et al., 2011). However, even slight variations in experimental stimuli can affect whether or not a

LAN is observed (see Molinaro et al., 2011, for a review). LANs tend to occur in conjunction with P600s, but P600s may be observed without an accompanying LAN.

The functional interpretation of the LAN, like that of the P600, remains contentious. Some authors suggest that the LAN may index working memory usage (Coulson et al., 1998; King & Kutas, 1995; Piai, Meyer, Schreuder & Bastiaansen, 2013), while others associate it with the detection of morphosyntactic errors (Friederici, 2002). Other authors have proposed that two functionally distinct LAN components exist, one corresponding to each of these interpretations (Fiebach, Schlesewsky, & Friederici, 2002). Some scholars have also noted that, since the topographical distribution and latency of the LAN, as well as its eliciting conditions, are all highly variable, it may be the case that various components described as LANs in the literature actually represent a broad range of functionally distinct components (Swaab et al., 2012). Alternatively, according to some views the LAN may not be an actual component, but an artefact caused by averaging ERP waveforms across populations of people who use different processing strategies (Tanner & Van Hell, 2014).

While there is much that remains debated about the N400, P600 and LAN, particularly regarding their functional interpretations, investigating these ERPs can still be highly informative when studying sentence processing. For the purposes of the current study, these ERPs will be examined primarily to determine whether the brain responds to certain phenomena in the same or different ways (i.e. whether the same or different ERPs are observed across conditions).

2.3. Speaker-identity integration in sentence processing

It is well accepted that speech processing mechanisms vary according to certain characteristics of the listener. For example, non-native listeners process certain aspects of speech differently to native listeners (e.g. Hoshino, Dussias, & Kroll, 2010), as do listeners with specific language impairment (SLI; e.g. Rispens & Been, 2007). Perhaps less obviously, however, the speech processing mechanisms of the listener have also been shown to vary as a function of the characteristics of the *speaker*.

When presented with speech, listeners are able to extract a large amount of information about the speaker from the speech signal, such as their sex, age, and so on. This information, rather than being discarded as non-speech-related, is rapidly integrated during speech processing, affecting how other aspects of the speech signal are processed (Nygaard, 2008). Thus, any given utterance could be processed in a variety of ways, potentially resulting in different interpretations, depending on the identity of the speaker by whom it was produced. This phenomenon has primarily been examined in relation to phonological and semantic processing.

The integration of speaker-identity information during phonological processing is of great importance, as accurate sound-to-phoneme mapping can only occur when phonetic variability between speakers is taken into account. This is generally termed ‘speaker normalisation’. The study of speaker normalisation examines how listeners are able to process potentially ambiguous phonetic information and map that information onto phonemes correctly by taking into account the identity of the speaker. The identification of vowels is a good example of this. Listeners are reliably able to classify vowels according to their first and second formant values (F1 and F2). However, F1 and F2 are ambiguous cues to vowel identity as the F1 and F2 values of adjacent vowels commonly overlap (Peterson & Barney, 1952). Ladefoged and Broadbent (1957) used synthesised speech to demonstrate that the perception of an ambiguous vowel can be shifted according to the pronunciation of other vowels by the same speaker, demonstrating that implicit information about the specific speaker who is producing the utterance is integrated online during phonological processing.

It has also been shown that the way in which phonetic input is mapped onto a phonological form can be manipulated by modifying the supposed characteristics of the speaker. For example, when listening to speech produced by an androgynous voice, listeners systematically varied their classification of the speaker’s vowels depending on whether they were informed that the speaker was male or female (Johnson et al., 1999). Similar results were found when groups of listeners were told that a speaker was either American or Canadian (Niedzielski, 1999). Furthermore, the categorical perception of fricatives has been shown to vary based on the perceived sex of the speaker (Munson, 2011). Thus there is strong evidence that listeners take the speaker’s identity into account when processing the phonetic information available in the speech stream and mapping that information onto phonemes.

Speaker-identity information at both the group and individual levels has also been shown to influence online semantic processing. A number of studies have investigated neural responses to sentences which are semantically congruous but are rendered incongruous by the group-level indexical information about the speaker conveyed through the speech signal, such as their age, sex or socio-economic status (e.g. Van Berkum et al., 2008). For example, the sentence *“I like to wear lipstick”* is semantically acceptable but generates a sex-based speaker incongruity when spoken by a male voice, considering that men are unlikely to wear lipstick.

Studies of this kind have shown differing neural responses between speaker-congruous and speaker-incongruous sentences using both EEG (e.g. Van Berkum et al., 2008) and fMRI (Tesink et al., 2009). However, the exact nature of the responses observed have varied. Some EEG studies reported an N400 response to speaker incongruities, much like semantic incongruities of other kinds (Van Berkum et al., 2008; Hanulikovà & Carreiras, 2015), while another observed a late positivity resembling a P600, subsequently labelled a late positive potential (LPP; Lattner & Friederici, 2003). Other studies have noted high levels of individual variation among participants, with some showing an N400, others an LPP and others with no response at all (Foucart et al., 2015, van den Brink et al., 2012). This variability has been attributed to the participants' language background (native versus late bilingual speakers; Foucart et al., 2015) or their level of empathy (van den Brink et al., 2012). There is also a possibility that the LPP is more often observed in response to gender-related speaker incongruities than speaker incongruities of other kinds (Lattner & Friederici, 2003; Van Berkum et al., 2008). Thus, although the exact mechanism and associated neural correlates remain unclear, it seems that speaker characteristics do influence the processing of semantic information online.

The influence of speaker-identity integration on semantic processing has also been examined in contexts where auditory (and sometimes visual) information identifying a *specific* speaker, rather than a class of speakers, is provided. For example, in an EEG study comparing neural responses to true or false statements made by a well-known politician, a media personality and an unknown speaker, the speakers' identifiability and their ability to act upon their statement were both found to influence how listeners responded to the utterance (Bornkessel-Schlesewsky, Krauspenhaar, & Schlesewsky, 2013). Similarly, in an eye-tracking task both adults and three- to ten-year-old children were shown to base their online predictions regarding upcoming words in a sentence on information provided about the speaker's identity (e.g. whether the speaker was a pirate or a princess; Borovsky & Creel, 2014). Thus online integration of speaker identity at both the group and individual levels have been shown to influence semantic processing.

However, in contrast to phonological and semantic processing, extremely few studies have investigated how speaker-identity integration may affect syntactic and/or morphosyntactic processing. The only existing study which has explored speaker-identity integration in syntactic or morphosyntactic processing using a neurophysiological measure was conducted by Hanulikovà and colleagues (2012). This study used EEG to investigate the effect of foreign-accent-related speaker-identity information on the processing of grammatical gender agreement in Dutch. Speaker identity was indeed found to modulate morphosyntactic processing, as listeners displayed a P600 in response to gender agreement violations produced by a native Dutch speaker, but not by a Turkish-accented speaker of Dutch. The authors hypothesised that this difference in neural responses may have been due to the influence of the listeners' expectations, based on the speakers' identity. Since gender agreement errors are commonly made by Turkish-accented Dutch speakers, the listeners may have used the speaker-identity information present in the accented speech signal to anticipate errors from the Turkish-accented speaker, and thus did not show a violation effect when these errors occurred. Thus the P600 was only observed in cases of speaker-incongruous grammatical violation and not when the violation was congruous with what would be expected of the speaker. There is therefore at least some evidence that, as with phonological and semantic processing, listeners do take into account the identity of the speaker when performing morphosyntactic processing.

2.4. Speaker-identity integration and foreign-accented speech

Previous studies of speaker-identity integration have manipulated a variety of speaker characteristics, including, for example, age, sex and socio-economic status (e.g. Van Berkum et al., 2008). The speaker characteristic manipulated in the current study (and in Hanulikovà and colleagues' (2012) study) is the speaker's foreign-accentedness. That is, whether the speaker has a foreign (in this case non-Australian-English) accent or not.

Foreign-accented speech is an ideal environment in which to examine speaker-identity integration, as accent has a fundamental relationship to speaker identity. The term 'accent' is typically defined with reference to identity, for example in the words of Crystal (2003, p. 458): an accent is composed of "features of pronunciation that signal regional or social identity". Similarly, Moyer (2013, p. 12) states that "accent is a reflection of our past experiences: languages known, regional and social upbringing, educational background, and affiliations with various speech communities and social networks", and that "accent is a medium through which we project individual style and signal our relationship to interlocutors" (p. 19). Thus when listeners perceive and integrate a speaker's accent, they are also, consciously or not, integrating information regarding the speaker's identity.

Foreign-accentedness also lends itself to investigating speaker-identity integration because it is a salient feature of the speech stream. Foreign-accented speech is generally distinguished from native speech by differences in phonetic realisation at the levels of segment, syllable and prosody (Major, 2001), meaning that many cues to foreign-accentedness are generally available. Listeners are able to utilise these cues to distinguish between native and non-native speakers of a language with a high degree of accuracy (Vieru, Boula de Mareüil, & Adda-Decker, 2010). Foreign-accentedness is thus a way in which identity is conveyed in the speech stream that is unlikely to go unnoticed by listeners.

2.5. Mandarin-accented English and subject-verb agreement

It is clear that further research investigating speaker-identity integration in speech processing is necessary, as very little is known about how syntactic and morphosyntactic processing may be affected by speaker identity. We do not yet know if Hanulikovà and colleagues' (2012) findings regarding speaker-identity integration in morphosyntactic processing generalise beyond Turkish-accented Dutch. Similarly, the question remains whether this effect will be observed in response to morphosyntactic errors other than gender agreement violations. Therefore, the current study aimed to investigate whether or not the findings of Hanulikovà and colleagues' could be generalised to the morphosyntactic processing of English. Due to the lack of grammatical gender agreement in English, it also aimed to investigate whether the effect would extend to subject-verb agreement violations, which perform a similar role to the previously-studied gender-agreement violations in signalling the relationships between sentence constituents. We investigated speaker-identity information in the guise of foreign-accented speech due to the reasons outlined in section 2.4, as well as to remain comparable to the study of Hanulikovà and collegues.

The particular foreign accent chosen for use in this study was Mandarin-accented English. This choice was made in part due to the widespread nature of this variety: 1.6% of Australian residents report speaking Mandarin at home, making Mandarin the most commonly-spoken language in Australia aside from English (Qpzm LocalStats Australia, 2016). The percentage of Mandarin speakers rises to 3.6% when considering only the greater Sydney area, rather than Australia as a whole (Ting & Walters, 2016). This gives Mandarin-accented English a similar social status and degree of familiarity to the Turkish-accented Dutch used in

Hanulikovà and colleagues' (2012) study, helping to make these studies more comparable. Additionally, as the effect observed in Hanulikovà and colleagues' study was thought to occur due to participants' prior experience with the accent in question, it was essential that, in investigating this phenomenon, an accent with which the participants were likely to be familiar was used.

Investigating the combination of subject-verb agreement violations and Mandarin-accented English was also motivated by an interesting interaction between these two phenomena. Morphologically marked subject-verb agreement errors in English occur in two types: omission and commission errors. Omission errors result from the omission of the third person singular -*s* in obligatory contexts, causing ungrammatical utterances such as **The father carefully cook*, as opposed to *The father carefully cooks*. Commission errors occur when the third person singular -*s* is included inappropriately, such as in **The fathers carefully cooks*, as opposed to *The fathers carefully cook*. While subject-verb agreement errors are common in Mandarin-accented English, there is an imbalance in the frequency of these error types, with over 90% of subject-verb agreement errors being omission errors (Jia & Fuse, 2007).

The imbalance in frequency of omission and commission errors in Mandarin-accented English provided an ideal scenario in which to examine the influence of speaker-related listener expectations. If Hanulikovà and colleagues were correct in their hypothesis that the lack of P600 in response to agreement errors produced by foreign-accented speakers was due to listeners expecting these errors to occur, we would predict in the current study that no P600 would be observed in response to omission subject-verb agreement errors, as these are likely to be expected by listeners, given their frequency. In contrast, we would predict that a P600 *would* be observed in response to commission errors, even when uttered by a foreign-accented speaker, because even though Mandarin-accented English speakers typically make subject-verb agreement errors, they are not usually of this type.

A potential alternative is that Hanulikovà and colleagues' results were not in fact due to listeners' expectations regarding foreign-accented speech, based on their prior experience, but were simply the result of a general effect of foreign-accentedness during speech processing. If this were the case, we would expect that no P600 would be observed in response to either omission or commission errors produced by the foreign-accented speaker in the current study.

These predictions assume, however, that the neural correlates of Dutch gender-agreement processing are shared by English subject-verb agreement processing. This is not necessarily the case. The LAN is not typically observed during Dutch gender-agreement processing (Hagoort & Brown, 1999), and was therefore not investigated in Hanulikovà and colleagues' study. However, the LAN is typically associated with aspects of morphosyntactic processing, including English subject-verb agreement (Osterhout & Mobley, 1995; Dube, Kung, Peter, Brock, & Demuth, 2016). The occurrence of the LAN in this context is not guaranteed, however: much variability has been observed.

A range of existing studies have examined the neural correlates of subject-verb agreement processing, reporting varying results. While the canonical response to subject-verb agreement errors is a biphasic LAN-P600 response (e.g. Dube et al., 2016; Jolsvai, Sussman, Csuhaj, & Csépe, 2011; see Molinaro et al., 2011, for a review), a number of studies have observed only a P600 (e.g. Tanner & Bulkes, 2015; Kos, Vosse, Van Den Brink, & Hagoort, 2010). In other studies, only an LAN (Kutas & Hillyard, 1983) or an N400 effect (Severens, Jansma, & Hartsuiker, 2008) was reported in an absence of a P600 effect. This variability in the components observed in response to subject-verb agreement violations highlights the fact that there is not yet a consensus on the neural correlates of subject-verb agreement processing.

An additional complicating factor in the current study was the use of foreign-accented speech, which has also been shown to affect ERP components. For example, the N400 component has been shown to have a smaller amplitude in response to foreign-accented speech than native speech (Goslin, Duffy, & Floccia, 2012). In other circumstances, however, topographical, rather than amplitude, differences have been observed (Hanulikovà et al., 2012). Additional neural resources are also proposed to be utilised in foreign-accented speech processing. fMRI studies have shown increased activation in the left superior temporal gyrus and superior temporal sulcus (Adank, Davis, & Hagoort, 2012) or bilaterally in the posterior superior temporal gyri and in the left inferior frontal gyrus (Adank, Noordzij, & Hagoort, 2012) when processing accented speech as compared to native speech (although both these studies used a novel accent, rather than a real foreign accent, which may have had an impact on the results; cf. Weber, Di Betta, & McQueen, 2014). The fact that foreign-accented speech is processed more slowly than native speech (Floccia, Butler, Goslin, & Ellis, 2009) and is associated with increased listening effort and greater use of executive processes such as working memory and cognitive control (Van Engen & Peelle, 2014) may also have an impact on the characteristics of the ERPs observed. Thus it is possible that other differences between ERP responses to native and foreign-accented speech may be observed, aside from the lack of P600 response predicted by Hanulikovà and colleagues' results.

2.6. Hypotheses of the current study

Taking this information into account, the following hypotheses were made regarding the current study. Firstly, we hypothesised that a P600, and possibly a LAN, would be observed in response to both omission and commission subject-verb agreement violations when produced by a native speaker of Australian English, in accordance with the existing findings described in section 2.5.

Secondly, we hypothesised that the neural responses observed in response to the subject-verb agreement errors produced by the foreign-accented speaker would differ from those observed in response to the native speaker. This could be realised in one of two ways. It may be the case that a P600 and possible LAN would not be observed in response to omission errors made by a Mandarin-accented speaker of English, but would be observed in response to commission errors by that speaker. This finding would reflect the difference in the degree of expectedness of these two error types and thereby provide evidence that speaker-identity integration is driven by listeners' expectations based on prior experience. Alternatively, it could be the case that the P600 and possible LAN would not be observed in response to either omission or commission errors made by the Mandarin-accented speaker, suggesting that listener expectations built on prior experience do not play a role, but instead that the observed effect is a general response to foreign-accented speech. The ERP components observed in response to the foreign-accented speaker could also differ from those in response to the native speaker in their latency and/or distribution, reflecting differences in the processing of foreign-accented speech as compared to native speech which were not specifically relevant to this study.

Thirdly, we hypothesised that individual differences would be observed between the responses of the participants, as in previous studies examining the influence of speaker-identity integration on semantic processing (Foucart et al., 2015, van den Brink et al., 2012, see section 2.3). This variation may potentially be attributed to participants' empathy levels (as in van den Brink et al., 2012), their level of exposure to foreign-accented English or Mandarin-accented English specifically, their working memory capacity (cf. Van Engen & Peelle, 2014), how accented the participant perceived the foreign-accented speaker's speech to be, or whether or not they were able to identify the foreign-accented speaker's accent. See Table 2.1 for a summary of hypotheses regarding the directionality of these effects.

Table 2.1 - Hypotheses regarding the potential effects of the between-subjects factors of interest included in the analysis of individual variation.

Between-subjects factor	Hypothesis
Empathy level	We hypothesise that more empathic listeners may show a smaller component amplitude in response to morphosyntactic violations produced by the foreign-accented speaker, as they are more able to cognitively empathise with this speaker and therefore have a stronger expectation that the speaker is likely to make errors. The more the errors are expected, the smaller the amplitude of the violation effect we would expect.
Exposure to foreign-accented English	We hypothesise that listeners with greater exposure to foreign-accented speech may show a smaller component amplitude in response to morphosyntactic violations produced by the foreign-accented speaker, as their familiarity with this type of speech may allow them to create stronger expectations regarding the presence of errors in Mandarin-accented speech.
Exposure to Mandarin-accented English	Alternatively, exposure to Mandarin-accented speech in particular may be required for listeners to show a smaller component amplitude in response to morphosyntactic violations produced by the foreign-accented speaker, as their familiarity with Mandarin-accented speech may allow them to create stronger expectations regarding the presence of errors.
Working memory capacity	As the processing of foreign-accented speech has been associated with increased use of working memory (Van Engen & Peelle, 2014), we hypothesise that listeners with greater working memory capacity may have a processing advantage in this task and thus be more sensitive to violations, showing violation effects with a greater amplitude.
Perception of speaker accentedness	We hypothesise that the more accented the listener perceives the foreign-accented speaker's speech to be, the more likely they will be to expect errors, thereby showing smaller amplitudes for the violation effect.
Identification of speaker accent	We hypothesise that if the listener is able to correctly identify the accent of the foreign-accented speaker, they will likely have strong expectations regarding the characteristics of that accent and therefore expect errors to occur, resulting in a smaller amplitude of the violation effect.

3. Methodology

In order to examine neural responses to subject-verb agreement violations in Mandarin-accented English, an EEG experiment was carried out following the method used by Hanulikovà, van Alphen, van Goch and Weber (2012). In the experiment, participants listened to sentences produced by a native Australian English speaker and a Mandarin-accented speaker of English. ERPs time-locked to subject-verb agreement errors (both omission and commission) were measured. Additionally, as in Hanulikovà and colleagues' study, ERPs associated with semantic violations were measured as a separate control condition. This was intended to confirm that the Mandarin-accented sentences were indeed being understood by the listeners, and to demonstrate that any lack of effect observed in response to the grammatical violations could not be attributed to listeners' inability to understand the speaker. ERP responses to the morphosyntactic-condition sentences were then correlated with a number of other variables to probe potential sources of individual variation in the results.

3.1. Participants

EEG data was collected from 30 monolingual native speakers of Australian English from the Sydney area (13M, 17F), whose parents were also native speakers of Australian English. The language background of the participants' parents was controlled in order to avoid effects of exposure to foreign-accented speech during language acquisition, which may influence speech perception later in life (Chen, Xu Rattanasone, Cox, & Demuth, in submission). The participants reported no hearing, language, neurological or cognitive impairments and were aged 18 to 33 years (mean = 22 years, s.d. = 5.3 years). All were right-handed according to the Edinburgh Handedness Inventory (Oldfield, 1971). Data from a further nine participants were excluded from analysis due to technical error (n = 4) or failure to meet the participation criteria[1] (n = 5). Some participants were undergraduate students from the Department of Psychology at Macquarie University and participated for course credit. Other participants were reimbursed $40 for their time. The Macquarie University Human Research Ethics Committee approved the experimental methods used in the study and written informed consent was obtained from all participants before beginning the test session.

[1] All participants completed an online screening questionnaire to ensure they met participation requirements before attending the session (see section 3.3). However, the excluded participants provided erroneous responses in their screening questionnaires which contradicted their responses in the pre-task questionnaire completed during the session.

3.2. Materials

For this experiment two groups of auditory stimuli were required: morphosyntactically correct and incorrect sentence pairs to investigate the phenomenon of interest (Table 3.1), and semantically correct and incorrect pairs for the control condition (Table 3.2). The morphosyntactic-condition sentences were of the form "The [noun] [adverb] [verb]", with the adverb in each case being either 'carefully', 'happily', or 'typically'. Each of these sentences occurred in a singular and plural subject version (25 of each, 50 in total). The semantic-condition sentences were of the form "The [noun] [verb] the [noun]" and contained a mixture of singular and plural subjects (25 in total).

Table 3.1 – Examples of morphosyntactic-condition sentences used in the task.

Grammatical singular	**Ungrammatical singular (omission errors)**	**Grammatical plural**	**Ungrammatical plural (commission errors)**
The father carefully cooks.	The father carefully cook.	The fathers carefully cook.	The fathers carefully cooks.
The sister happily shops.	The sister happily shop.	The sisters happily shop.	The sisters happily shops.
The pilot typically sits.	The pilot typically sit.	The pilots typically sit.	The pilots typically sits.
The lady carefully knits.	The lady carefully knit.	The ladies carefully knit.	The ladies carefully knits.
The kitten happily licks.	The kitten happily lick.	The kittens happily lick.	The kittens happily licks.
The tiger typically bites.	The tiger typically bite.	The tigers typically bite.	The tigers typically bites.

Table 3.2 – Examples of semantic-condition sentences used in the task.

Semantically correct	**Semantic violation**
The lady bites the cupcake.	The lady bites the suitcase.
The barman cools the cocktail.	The barman cools the carpet.
The helper cuts the carrot.	The helper cuts the buses.
The mother feeds the baby.	The mother feeds the doorbell.
The captain leads the soldiers.	The captain leads the teacup.
The actor leaves the party.	The actor leaves the ceiling.

In both the morphosyntactic- and semantic-condition sentences, all verbs were CVC monosyllables and were in present tense, while all nouns were disyllabic and had initial stress to create a consistent prosodic contour across sentences. All nouns and verbs had a log frequency greater than 3.0 in the SUBTLEX-UK database (Van Heuven, Mandera, Keuleers, & Brysbaert, 2014). The nouns and verbs did not contain any consonant clusters (except those created by the addition of the third person singular -*s*). This was to ensure the target words were maximally intelligible when produced by the foreign-accented speaker, as clusters, particularly in coda position, are difficult for speakers of Mandarin-accented English to produce (Xu & Demuth, 2012). All subject nouns were animate (the canonical animacy class for subjects), as animacy violations have been shown to increase the amplitude of observed P600 effects (Szewczyk & Schriefers, 2010), and consistent thematic roles were maintained across sentences (all subjects were agents, all objects were patients, however animacy was not controlled for the object nouns).

In the morphosyntactic-condition sentences only, all verb stems ended with voiceless stops. Stem-final stops were chosen to facilitate time-locking during the EEG experiment, and voicelessness was necessary in order to ensure that all instances of the third person singular -*s* were produced with the same [s] allophone, rather than [z], which occurs after voiced segments. In the semantic sentences, all object nouns were selected to begin with a stop, affricate or /s/ to facilitate splicing.

In addition to the morphosyntactic- and semantic-condition sentences, 100 filler sentences were created. Each filler sentence was either six or seven syllables long to be of comparable length to the sentences in other conditions. Consonant clusters were allowed in these sentences as maximal intelligibility was less of a concern. Because of this, the fillers produced by the foreign-accented speaker were expected to sound more strongly accented than the target sentences, reinforcing the speaker's foreign-accentedness. The fillers were divided equally between the two speakers.

Two female speakers, both undergraduates at Macquarie University, were recruited to record the sentences. The native speaker was 24 years old and was a monolingual speaker of Australian English. The Mandarin-accented speaker was 20 years old and was selected from among three candidates as being the speaker with the most canonical yet intelligible Mandarin accent. She was a native speaker of Mandarin from Beijing who had been living in Australia for two years. Both speakers had previously recorded auditory stimuli for other studies and so were experienced in the recording procedure.

The two speakers recorded the correct sentences and fillers only. Sentences with violations were later created by splicing the correct sentences to prevent the appearance of phonetic cues to ungrammaticality or semantic unacceptability occurring before the overt violation (cf. Hasting & Kotz, 2008; Royle, Drury & Steinhauer, 2013). Sentences were elicited from the native speaker using a question and answer method to maintain consistent prosody and to ensure that the final word of each sentence was focused. This method was too challenging for the foreign-accented speaker, so the foreign-accented sentences were elicited by having the speaker imitate a native Australian English speaker. A comparable speech rate was maintained throughout the recordings and between speakers.

The correct sentences were then spliced using Praat software (Version 5.4.08; Boersma & Weenink, 2015) to create the necessary morphosyntactic and semantic violations. Subject-verb agreement violations were created by splicing the adverb and verb from the plural sentences onto the singular sentences to create omission errors and splicing the adverb and verb from the singular sentences onto the plural sentences to create commission violations (Figure 3.1). Splicing occurred at the initial stop in ‘carefully’ or ‘typically’, or at the medial /p/ in ‘happily’.

Semantic violations were created by cross-splicing the object noun phrases between sentences. It was initially planned that semantic sentences would be spliced at the onset of the object noun, however the resulting sentences sounded highly unnatural. The sentences were therefore spliced at the beginning of the object noun phrase (i.e. at the onset of ‘the’) instead, rendering the control of noun-initial consonants described above unnecessary (Figure 3.1). With splicing, 400 sentences were created in total (Table 3.3; Appendix A).

Figure 3.1 – Diagram showing splicing locations for sentences containing morphosyntactic and semantic violations.

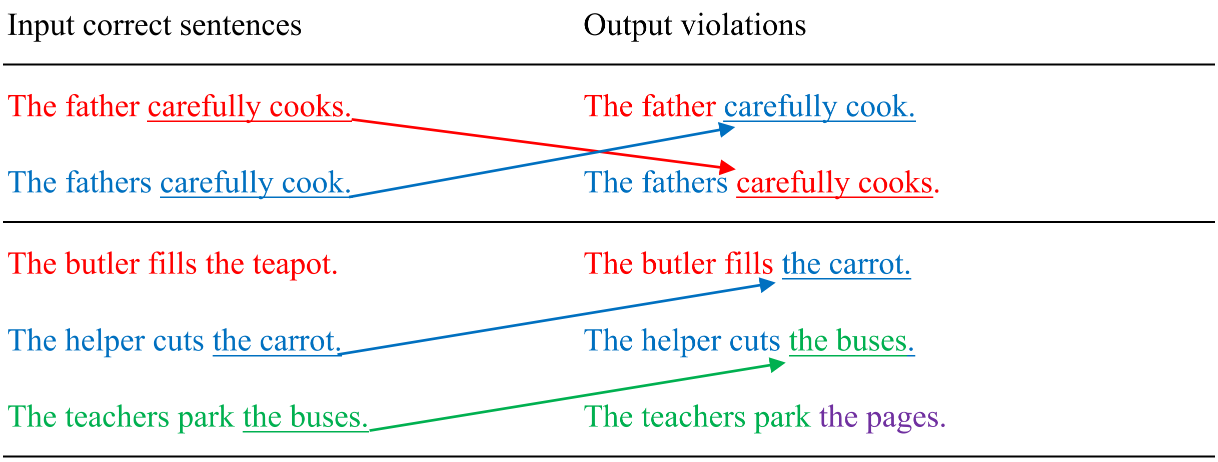


Table 3.3 – Number of sentences in each condition produced by each speaker.

			Native	Foreign-accented
Morphosyntactic	Correct	Singular	25	25
		Plural	25	25
	Violation	Singular (omission)	25	25
		Plural (commission)	25	25
Semantic		Correct	25	25
		Violation	25	25
Fillers			50	50
Total			**200**	**200**

3.2.1. Validation of materials

A listening task was carried out in order to ensure the stimuli were suitable for use in the EEG task. Thirteen undergraduate students (13F) from the Department of Linguistics at Macquarie University participated for course credit. All participants were monolingual native speakers of Australian English aged 19 to 23 years (mean = 20 years, s.d. = 1.3 years) and had parents who were also native speakers of Australian English, to match the population recruited for the EEG task. They reported no hearing, language, neurological or cognitive impairments.

The listening task had three components. Firstly, participants listened to all of the correct target sentences (i.e. both semantically and morphosyntactically correct) produced by the Mandarin-accented speaker and rated each for intelligibility on a seven-point scale ('Impossible to understand' (1) to 'Very easy to understand' (7)). A mean rating greater than 4.5 was used to signify that the given sentence was sufficiently intelligible to be included in the EEG task.

Secondly, participants rated all sentences produced by the native speaker in the semantic condition (both correct and incorrect) on a seven-point scale for plausibility ('Makes no sense' (1) to 'Makes perfect sense' (7)). In this case, correct sentences with mean ratings greater than five and incorrect sentences with mean ratings less than three were considered sufficiently plausible and implausible respectively to be included.

Finally, to ensure that the foreign-accented sentences were actually perceived as being foreign-accented, participants were asked at the end of the task to rate the degree of accentedness of the foreign-accented speaker on a seven-point scale ('No foreign accent' (1) to 'Very strong foreign accent' (7)). As this accentedness rating was based on target sentences only, participants then listened to an additional ten consecutive filler sentences produced by the foreign-accented speaker and made a second accentedness rating based on these sentences. This second rating was included so that a mean accentedness rating could be calculated for all sentences heard during the task, as, due to the additional phonological constraints placed on the target sentences, it could be the case that target sentences and fillers would be perceived as having differing levels of accentedness.

The results of the listening task revealed that 22 of the 75 sentences rated for intelligibility (29%) were insufficiently intelligible. Additionally, one semantically correct sentence and its incorrect counterpart were not sufficiently plausible and implausible respectively. These sentences were discarded and replacement sentences were recorded using the same procedure as outlined above.

These new sentences were rated for intelligibility and plausibility in the same manner as previously. The participants in this case were 10 monolingual native speakers of Australian English (10F) aged 20 to 50 years (mean = 32 years, s.d. = 10.3 years) whose parents were also native speakers of Australian English, as in the previous group. They reported no hearing, language, neurological or cognitive impairments. The results of this second listening task yielded sufficient sentences with acceptable ratings to replace those discarded from the initial stimuli set.

The accentedness of the foreign-accented speaker received a mean rating of 4.75 (s.d. = 0.91) based on the target sentences[2] and 5.39 (s.d. = 0.99) based on the fillers across the two listening tasks. There were no significant differences between the ratings given in the two tasks, according to Wilcoxon rank sum tests (p = .932 for target sentences, p = .531 for fillers). The difference between the ratings given to the target and filler sentences was

[2] This rating is based on results from 20 of the 23 participants, as the sentence presentation program initially did not allow participants enough time to respond to this question before moving on automatically. This was amended after the first three participants.

significant (p = .001) according to a paired samples Wilcoxon rank sum test[3], demonstrating that the fillers were indeed rated as being more strongly accented than the target sentences. The sentences received a mean accentedness rating of 5.09 (s.d. = 1.00) across both sentence types in the two listening tasks, suggesting that the speaker's accent was perceived as being moderate to strong overall.

3.3. Procedure

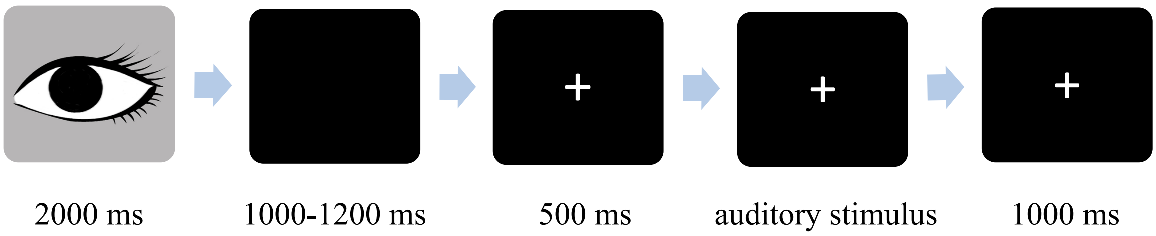
Before attending the test session, all participants completed an online screening questionnaire to confirm that they fulfilled the participation criteria for the experiment (Appendix B). Upon arriving for the session, participants first completed an auditory digit span task to assess their working memory capacity (Wechsler, Coalson, & Raiford, 2008). They then filled in a pre-task questionnaire (Appendix C) during EEG setup. The questionnaire was designed as follows: questions one to five were intended to confirm that the participant fulfilled the criteria to participate in the study (confirming the results of the screening questionnaire) and questions six to 11 collected information regarding potential factors which may influence the participant's neural responses, to be used in the analysis of individual variability. These factors included education level, exposure to foreign-accented speech and a measure of cognitive and affective empathy (the Empathy Quotient Short Form; Wakabayashi et al., 2006).

EEG data was then collected while participants sat in a comfortable chair in front of a computer monitor in a sound-attenuated, electromagnetically shielded room. During the task participants were presented with one of two lists of the 400 stimuli sentences via in-ear headphones. The sentences were presented in pseudorandomised order across eight blocks of 50 sentences each, with a practice block of ten sentences at the beginning of the experiment to familiarise participants with the task procedure. Each trial began with a cartoon image of an eye, presented on the screen for 2000 ms. Participants were instructed to blink whenever this image appeared on the screen. This was followed by a blank screen randomly varying in duration between 1000 and 1200 ms. A fixation cross then appeared on the screen for 500 ms before the onset of the auditory stimulus sentence and remained on the screen until 1000 ms after the completion of the sentence (Figure 3.2).

[3] The Wilcoxon rank sum test was conducted using the ratings from only 20 of the 23 participants (see previous footnote).

Participants were presented with five comprehension questions per block at irregular intervals to ensure they were attending to the sentences (40 questions in total). Questions always referred to the sentence immediately before. Responses to these questions were recorded using a button box. All participants performed at or near ceiling for comprehension accuracy, with a mean score of 39.5 out of 40 (98.7%). No participant scored below 37 out of 40 (92.5%). Therefore no participants were excluded for failing to attend to the task. Additionally, a short game of 'Snake' controlled by the button box appeared five times per block for participants to play in order to maintain their attention throughout the experiment.

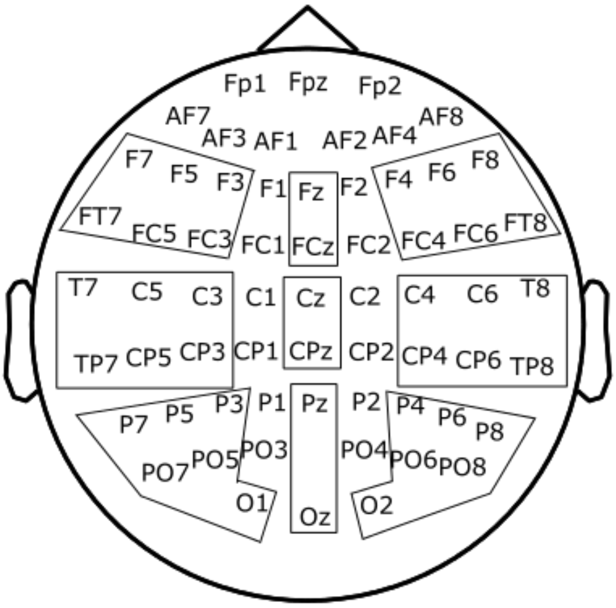
Figure 3.2 – Schematic diagram of the trial structure.



Continuous EEG data was recorded using Curry software (Version 7; Compumedics Ltd., USA) from 64 Ag/AgCl scalp electrodes mounted on an electrode cap (Easycap, Brainworks, GmbH) according to the International 10–20 system (Jasper 1958; Fpz, Fz, FCz, Cz, CPz, Pz, Oz, Fp1/2, AF7/3/1/2/4/8, F7/5/3/1/2/4/6/8, FT7/8, FC5/3/1/2/4/6, T7/8, C5/3/1/2/4/6, M1/2, TP7/8, CP5/3/1/2/4/6, P7/5/3/1/2/4/6/8, PO7/5/3/4/6/8, O1/2; Figure 3.3). Ground was located at AFz. Additional electrodes were placed above and below the right orbit and at the outer canthus of each eye to monitor electrooculographic activity. Electrode impedances were kept below 10kΩ. Electrical activity was recorded from both mastoids. The left mastoid was used as the on-line reference. The EEG signal was digitised at a sampling rate of 1000Hz and filtered with a .05-100Hz bandpass filter using a Neuroscan SynAmps2 DC Amplifier (Compumedics Ltd., USA).

The EEG recording lasted between one and one and a half hours for each participant. After the completion of the EEG recording, participants were given a post-task questionnaire (Appendix D) in which they were asked to identify the nationality of the foreign-accented speaker and rate the strength of their accent on a seven-point scale from 'No foreign accent' (1) to 'Very strong foreign accent' (7). In all, the testing session for each participant lasted approximately two and a half hours.

Figure 3.3 – Electrode layout, including regions of interest for statistical analysis.



3.4. Data processing

EEG data was processed in Matlab (Version R2015a; MathWorks, Massachussets, U.S.A) using the Fieldtrip toolbox (Version 20160719; Oostenveld, Fries, Maris, & Schoffelen, 2010). Epochs of approximately 5500 ms were extracted from the continuous data stream (from 1000 ms before the presentation of the fixation cross to 1500 ms after the offset of the audio file for each trial). An Independent Component Analysis (ICA) was then performed on the data to remove artefacts caused by eye-blinks and saccades, and, for some participants, their pulse, which appeared in the EEG signal. For the ICA, trials with extreme values were first rejected so as to make the input as consistent as possible. Then the ICA was performed and all components resembling blinks, saccades and pulses were identified. These components were then mathematically removed from the data and the signal was back-projected onto the complete dataset (including trials which were initially rejected).

After ICA, the data was re-referenced to the average of the mastoids. A baseline correction was then performed. For the morphosyntactic-condition sentences the baseline period was the 200 ms interval before the beginning of the trial (i.e. when the participants were viewing a blank screen). For the semantic-condition sentences the baseline period was the 200 ms interval immediately preceding the time-locking point of the ERP. Different baseline periods were chosen based on the results of a cluster-based permutation test performed on data with no baseline correction (Appendix E; see section 3.5 for a description of cluster-based permutation tests). These tests revealed that, in the semantic condition only, the waveforms elicited by correct and incorrect sentences differed significantly during the proposed pre-trial baseline period. The use of different baseline periods for the morphosyntactic-condition and semantic-condition sentences was considered acceptable as these conditions were treated separately throughout data analysis and were never compared to each other.

A zero-phase fourth-order Butterworth IIR low-pass filter of 35Hz was applied to the dataset, with a 12dB per octave slope and a 6dB cut-off. Epochs were then redefined as the period from the beginning of the baseline period to 1500 ms after the time-locking point (i.e. 1500 ms after the release burst of the verb-final stop in morphosyntactic-condition sentences and 1500 ms after the onset of the object noun in semantic-condition sentences).

Trials with artefacts were then removed from the data set. Trials with artefacts were defined as those epochs in which the amplitude variance was greater than 1000µV and/or the maximum z value was greater than 4. In total, 1398 trials were rejected (15.5% of all trials). The distribution of rejected trials across conditions is shown in Table 3.4. There was a notable difference in the number of trials rejected between morphosyntactic (19.5% rejected)[4] and semantic (7.6% rejected) conditions. This was presumably due to the difference in epoch length between these sentence types, caused by the differing locations of the baseline period. Again, this was not deemed problematic as morphosyntactic- and semantic-condition sentences were never compared during statistical analysis. Within the morphosyntactic and semantic conditions, no notable difference in number of rejections was observed between sentence types. Once trials containing artefacts had been rejected, the data was averaged within sentence types for each participant, then a grand average for each sentence type was calculated across participants.

[4] While the relatively high number of rejections among the morphosyntactic-condition trials resulted in a suboptimal number of remaining trials per condition, this was considered to be unavoidable as increasing the number of trials per condition would have resulted in an unfeasibly long EEG task in which participants would struggle to maintain attention.

Table 3.4 – Number of trials rejected across all participants.

			Native	Foreign-accented
Morphosyntactic	Correct	Singular	133 (17.7%)	156 (20.8%)
		Plural	159 (21.2%)	141 (18.8%)
	Violation	Singular (omission)	142 (18.9%)	141 (18.8%)
		Plural (commission)	142 (18.9%)	157 (20.9%)
Semantic		Correct	62 (8.3%)	55 (7.3%)
		Violation	56 (7.5%)	54 (7.2%)

3.5. Statistical analysis

In this study the morphosyntactic-violation sentences were the main focus and the semantic-violation sentences were included purely to validate the results elicited by the morphosyntactic violations. The morphosyntactic and semantic data were therefore treated separately throughout statistical analysis.

In order to determine the time windows to be used for further statistical analysis, cluster-based permutation tests were conducted to determine the time points at which the grand average waveforms of corresponding correct and incorrect conditions deviated from each other (Maris & Oostenveld, 2007). To conduct a cluster-based permutation test, a t-test is first performed on the amplitude values of the two waveforms at each individual sampling point. Sampling points at which the difference between the waveforms is found to be significant ($p < 0.05$, two-tailed) are then formed into clusters based on both temporal and spatial adjacency, and the polarity of the effect (i.e. t-values associated with sampling points in a single cluster must be either all positive or all negative). Cluster-level statistics are then calculated by adding together all t-values within the cluster. The largest of these cluster-level statistics are then tested for significance against clusters of samples selected using 1000 random partitions.

In the current study, cluster-based permutation tests were performed to compare the grand average waveforms for the correct and incorrect variants of each sentence type: semantically correct versus semantic violation, morphosyntactically correct plurals versus omission errors and morphosyntactically correct singulars versus commission errors for each of the two speakers. The choice was made to compare correct plural-subject sentences and omission error sentences (which have singular subjects) because the third-person singular *-s* morpheme is not present in either of these sentence types. Similarly, correct singular-subject and commission error sentences were compared as the third-person singular *-s* was present in both (cf. Dube et al., 2016). Thus the presence of the morpheme was not changed between compared conditions, but the context in which the morpheme occurred was. This prevented the difference in acoustic information presented after the time-locking point (i.e. /p, t, k/ in the plural correct and omission error sentences and /ps, ts, ks/ in the singular correct and commission error sentences) from acting as a confound (see Steinhauer & Drury, 2012, for further discussion of target versus context manipulation).

The beginning and end points of the significant clusters identified by the cluster-based permutation tests (rounded to the nearest 10 ms) were subsequently used to define the time windows for analysis in the corresponding conditions. Therefore, the mean amplitude of the ERP response during the identified time windows for each participant in each condition for each region of interest (ROI) was extracted. Nine ROIs were defined, corresponding to left, midline and right electrodes located in the anterior, central and posterior regions. For each midline ROI, the mean of two representative electrodes was taken, while the mean of six representative electrodes was taken for each lateral ROI (Figure 3.3; anterior midline: Fz, FCz; central midline: Cz, CPz; posterior midline: Pz, Oz; left anterior: F3/5/7, FC3/5, FT7; right anterior: F4/6/8, FC4/6, FT8; left central: C3/5, T7, CP3/5, TP7; right central: C4/6, T8, CP4/6, TP8; left posterior: P3/5/7, PO5/7, O1; right posterior: P4/6/8, PO6/8, O2).

Two omnibus ANOVAs were then conducted on the extracted mean amplitude values, one for the semantic-condition data and one for the morphosyntactic-condition data. For the semantic-condition data, the ANOVA had a 2x2x3x3 design: Correctness (correct or incorrect) by Speaker (native or foreign-accented) by Anteriority (anterior, central or posterior) by Laterality (left, midline or right). Correctness and Speaker were considered to be the primary factors of interest, and thus post hoc analyses were performed interactions involving these factors only. For the morphosyntactic-condition data, the ANOVA had a 2x2x2x3x3 design: Grammaticality (grammatical or ungrammatical) by Speaker (native or foreign-accented) by Presence of *-s* (*-s* or no *-s*) by Anteriority (anterior, central or posterior) by Laterality (left, midline or right). Grammaticality, Speaker and Presence of *-s* were the main factors of interest for the morphosyntactic-condition data. Thus post hoc tests were performed on only the interactions involving these factors.

3.6. Individual differences

In order to investigate factors which may have been responsible for individual variation in the morphosyntactic-condition data, a linear regression model was used. The main factors of interest were the participant's exposure to foreign-accented speech, exposure to Mandarin-accented speech specifically, working memory capacity, empathy level, perception of how accented the foreign-accented speaker's speech was, and whether or not they were able to correctly identify the speaker's accent. The indices used for these factors were as follows.

Exposure to foreign-accented and Mandarin-accented speech was calculated from the participant's responses to questions seven to ten of the pre-task questionnaire (Appendix C). In these questions, participants listed locations where they had lived, both inside and outside Australia, and gave frequency ratings for each foreign (i.e. non-Australian-English) accent they heard regularly on a five-point scale (daily, weekly, monthly, rarely (i.e. less than once per month), never). The frequency rating for each accent was transformed into an exposure score for that accent (daily = 4, weekly = 3, monthly = 2, rarely = 1, never = 0). Additionally, if participants reported living outside Australia at some point during their lives, the exposure score for the accent corresponding to the variety spoken in that country received an additional 5 points due to the exposure received from multiple sources, rather than a single speaker. For places of residence within Australia, a further 1 point was added to the exposure score of the relevant accent for each 10% of the population of the suburb that speaks that language at home, according to Australian Bureau of Statistics data (Qpzm LocalStats Australia, 2016). Having calculated the exposure score for each accent listed by the participant, the participant's exposure score for Mandarin alone was used as the index for that participant's level of Mandarin exposure. For the measure of exposure to foreign-accented speech in general, the sum of the exposure scores for all reported accents was used.

Scores from the auditory digit span task were converted to an index of working memory according to the Wechsler Adult Intelligence Scale (Wechsler et al., 2008). Scores calculated from the Empathy Quotient Short Form according to the method laid out by Wakabayashi and colleagues (2006) were used to provide an index of participant empathy. Participants' responses to question one of the post-task questionnaire (Appendix D) were coded as being either correct or incorrect and were used as a measure of accent identification accuracy.

Participants' responses to question two of the post-task questionnaire were used to represent how strong the participants perceived the Mandarin-accented speaker's speech to be.

The dependent variable of the linear regression model was the mean amplitude of the ERP component observed during the time period defined by the cluster-based permutation test. The model was designed to include Speaker, Violation type, Age, Gender, Education, Digit span score, Empathy Quotient score, Mandarin-accent exposure score, Foreign-accent exposure score, Accentedness rating score and Accent identification accuracy as predictors. Interaction terms were not included to facilitate model interpretation, given the large number of factors. In order to conduct the linear regression analysis, the *step* model-selection function was used in R (R Core Team, 2014). For each of the three linear regression models, this function provided the best fitting model based on a backward stepwise algorithm.

4. Results

4.1. Morphosyntactic violations

Neural responses to both commission and omission subject-verb agreement errors produced by a native speaker of Australian English and a Mandarin-accented speaker of English were analysed in this study in order to determine whether neural responses to the errors differed between speakers, and whether the difference was modulated by the error type.

Visual inspection of the grand average waveforms for each sentence type revealed no consistent patterns (Figure 4.1). However, cluster-based permutation tests carried out on the morphosyntactic-condition data showed a significant difference between the responses observed to sentences containing commission errors produced by the native speaker and their grammatical counterparts. This was represented by a significant negative cluster across the central and posterior electrode regions lasting from 207 to 326 ms post-stimulus-onset (i.e. post verb-final-stop burst release, $p = 0.033$). No significant differences were found between any of the remaining violation conditions (omission errors produced by the native speaker and both commission and omission errors produced by the foreign-accented speaker) and their grammatical counterparts (see Appendix E for details of further analyses which were carried out).

Figure 4.1 – Waveform and topography plots for both native-speaker and foreign-accented speaker morphosyntactic-condition responses.

Waveforms (native speaker): Commission-error condition shown in blue, omission-error condition in red. Responses to grammatical sentences shown as solid lines, responses to violations as broken lines. Negativity plotted downwards. Time window for analysis shown as a grey box. Note that the 200 ms pre-trigger period shown here is not the baseline period.

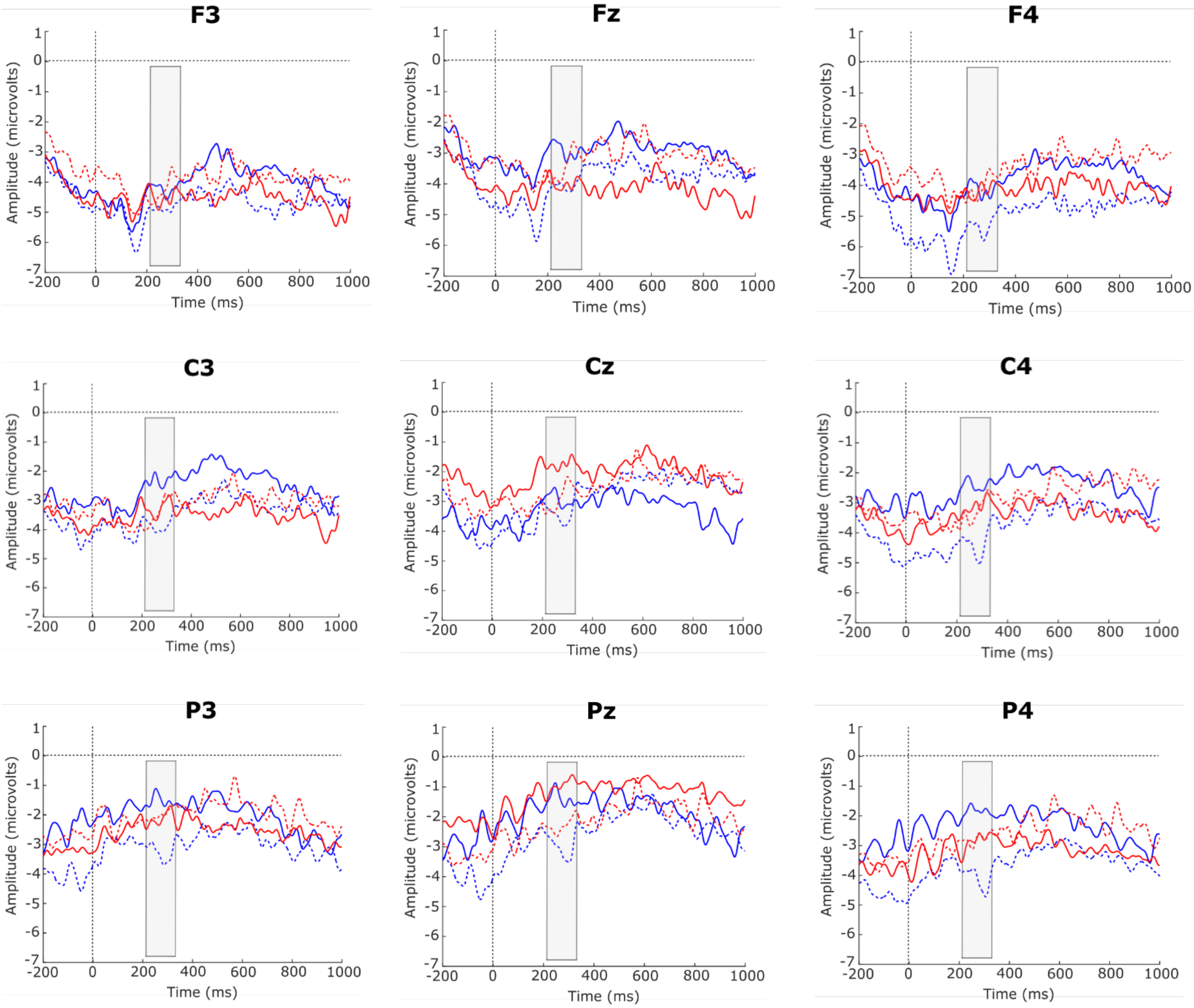


Figure 4.1 cont'd

Waveforms (foreign-accented speaker): Commission-error condition shown in blue, omission-error condition in red. Responses to grammatical sentences shown as solid lines, responses to violations as broken lines. Negativity plotted downwards. Time window for analysis shown as a grey box. Note that the 200 ms pre-trigger period shown here is not the baseline period.

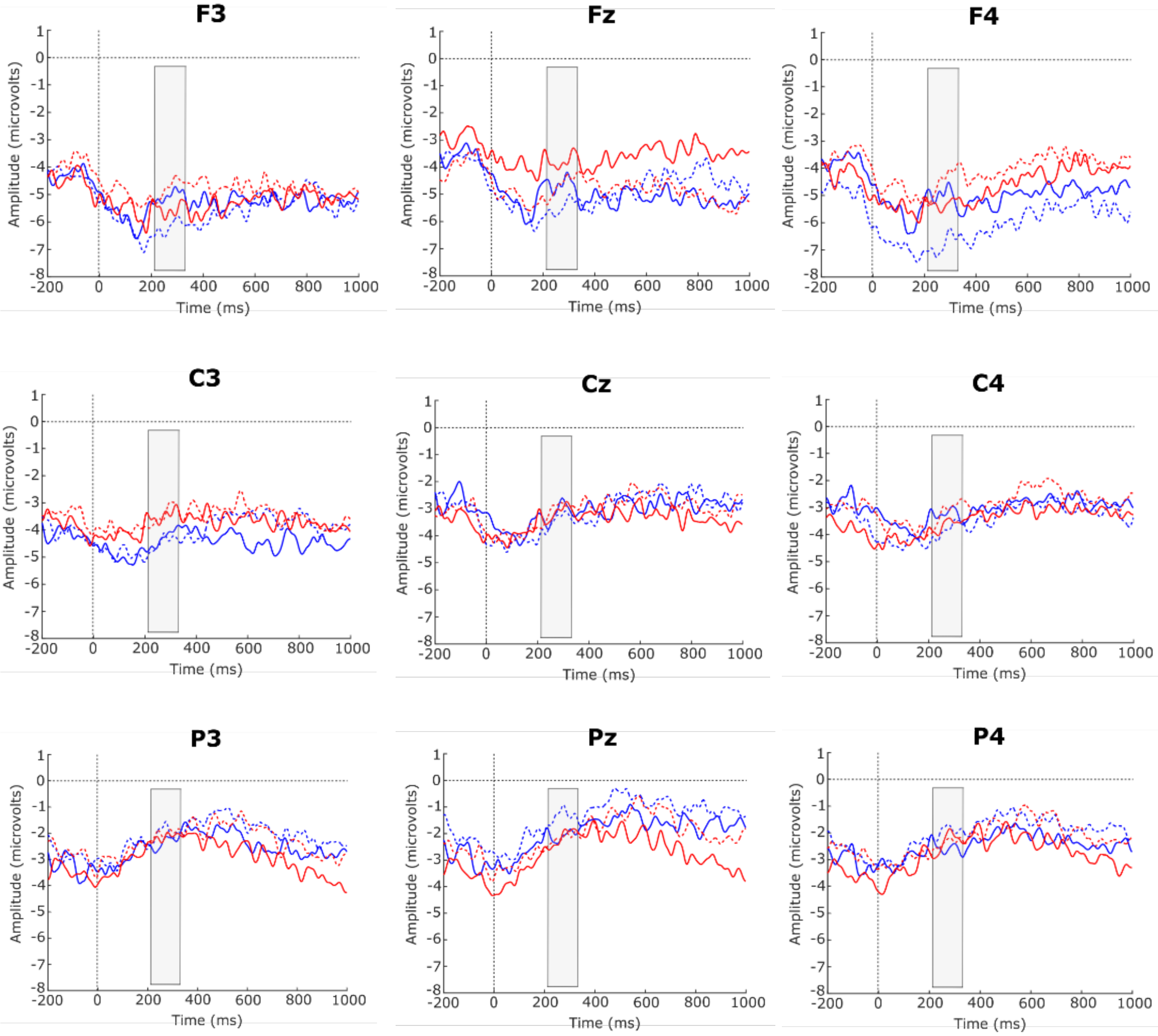
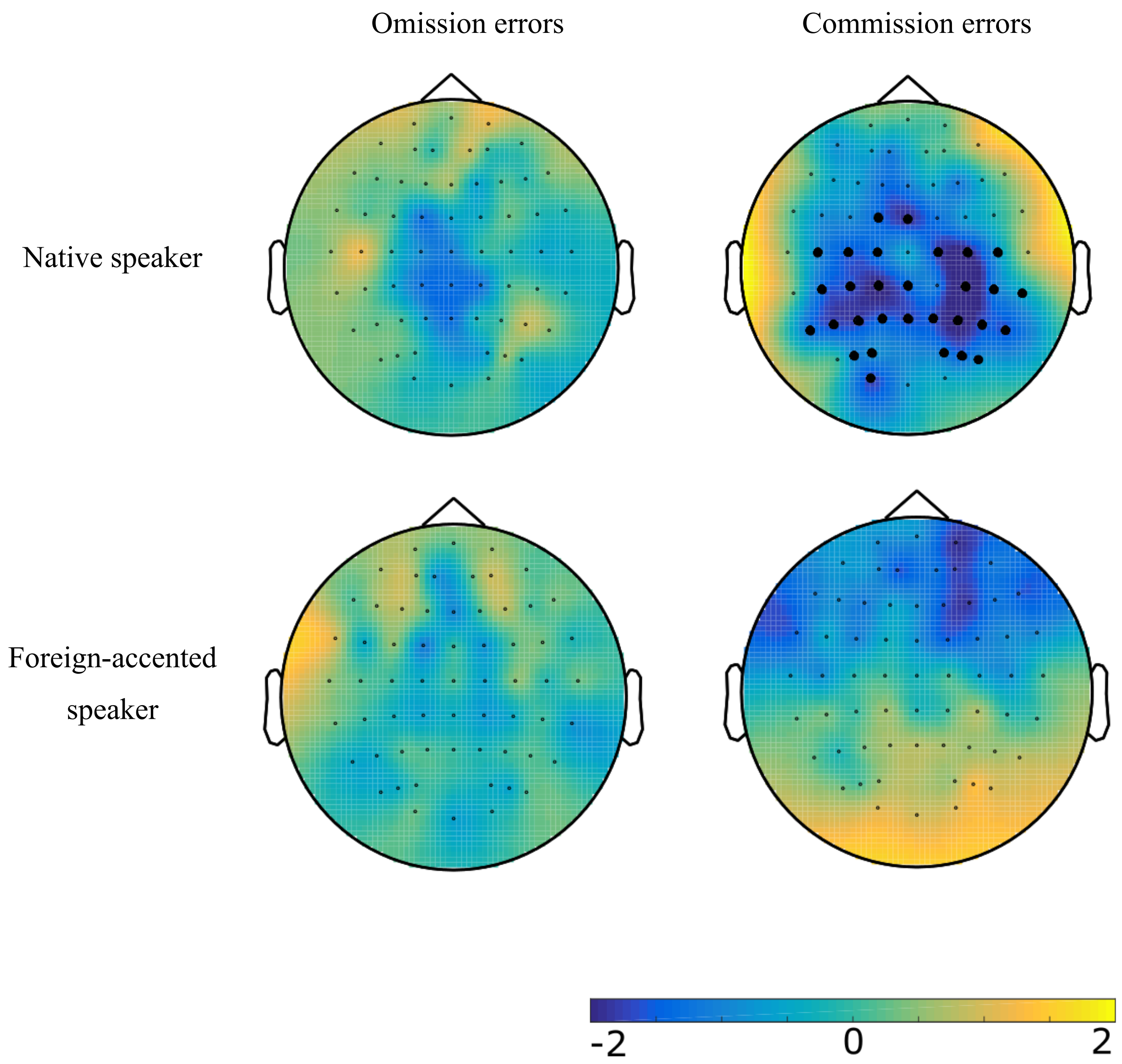


Figure 4.1 cont'd

Topography plots: Electrodes contributing to the significant cluster are shown as filled black circles.



The mean amplitude of the response in each scalp region during the time window identified by the cluster-based permutation test (210-330 ms) was then entered into an omnibus ANOVA to further examine the effects of the key factors of Grammaticality, Speaker and Presence of *-s* (Greenhouse-Geisser corrected values are reported here where necessary). While there were no significant main effects of any of these factors, a significant interaction of Speaker by Anteriority was found ($F[1.35, 39.11] = 18.35$, $p < 0.001$, $\eta^2 = 0.008$), as well as a significant five-way interaction of Grammaticality by Speaker by Presence of *-s* by Anteriority by Laterality (F[2.88, 83.61] = 3.67, p = 0.017, η^2 = 0.001).

To facilitate interpretation of these interactions, post hoc tests were conducted. For the interaction between Speaker and Anteriority, three pairwise t-tests (with a Bonferroni-corrected alpha of 0.017) were conducted to compare responses to foreign-accented and native-speaker sentences for each of the three levels of Anteriority (that is, anterior, central and posterior electrodes). These comparisons revealed that significantly more negative neural responses to foreign-accented sentences in general were only observed at the anterior electrodes (t[29] = 4.42, p < 0.001, 98.3% CI = [0.37, 1.35], d = 0.81).

To examine the Grammaticality by Speaker by Presence of *-s* by Anteriority by Laterality interaction, the data was split by Anteriority and a separate ANOVA was performed for each level of Anteriority (anterior, central and posterior). The ANOVA for central electrode sites showed no significant results. The ANOVA for anterior electrode sites showed a main effect of Speaker (F[1, 29] = 19.53, p < 0.001, η^2 = 0.028) and a significant interaction of Grammaticality by Presence of *-s* (F[1, 29] = 5.23, p = 0.029, η^2 = 0.009). The Grammaticality by Presence of *-s* interaction at the anterior electrodes was further investigated using two pairwise t-tests comparing responses to grammatical and ungrammatical sentences for each level of Presence of *-s* (Bonferroni-corrected alpha was 0.025). Neural responses to ungrammatical sentences at the anterior sites were found to be significantly more negative only when an *-s* was present (i.e. the violation was a commission error; t[29] = 2.70, p = 0.011, 97.5% CI = [0.11, 1.65], d = 0.49).

The ANOVA conducted on responses at the posterior electrode sites showed a three-way interaction of Grammaticality by Speaker by Presence of *-s* (F[1, 29] = 10.45, p = 0.003, η^2 = 0.016). To interpret this interaction, the data was once again split, this time according to Speaker. Two ANOVAs were performed, one on responses to native and one on responses to foreign-accented speech. The responses to sentences produced by the native speaker showed a significant Grammaticality by Presence of *-s* interaction (F[1, 29] = 4.49, p = 0.043, η^2 = 0.021). Pairwise t-tests were conducted, comparing grammatical and ungrammatical sentences for each value of Presence of *-s* (Bonferroni-corrected alpha was 0.025). These revealed that, for responses to the native speaker at the posterior electrode sites, responses to ungrammatical sentences were more negative than for grammatical sentences only when an *-s* was present (i.e. the violation was a commission error; t[29] = 2.81, p = 0.009, 97.5% CI =

[0.20, 2.34], $d = 0.51$). For sentences produced by the foreign-accented speaker, a significant interaction of Laterality by Presence of *-s* was observed at the posterior electrodes ($F[2, 58] = 3.30$, $p = 0.044$, $\eta^2 = 0.004$). In this case, further pairwise t-tests were not carried out, as determining the laterality of a general effect of Presence of *-s* that did not interact with Grammaticality was not considered relevant to achieving the aims of the study.

To summarise the results of the analysis of morphosyntactic-condition data, cluster-based permutation tests revealed a significant difference between responses to grammatical and ungrammatical sentences only the time window from 210 to 330 ms post-stimulus-onset, and only for the commission errors produced by the native speaker. Therefore, data from this time window across all conditions was entered into an omnibus ANOVA. The results of this ANOVA and the associated post hoc tests revealed that neural responses in the 210 to 330 ms time window did indeed show significant effects associated with the key variables of Grammaticality, Speaker and Presence of *-s.* A negativity associated with foreign-accented sentences in general was observed at the anterior electrodes, which was exacerbated by the presence of a commission error. Additionally, a broadly-distributed negativity (observed at the anterior and posterior electrodes) was observed for commission errors produced by the native speaker[5].

4.2. Individual differences

In order to explore possible factors which may be responsible for individual variation in the amplitude of the components observed, linear regression models were conducted. While it was initially planned to only conduct one model for the component observed in each time window, the fact that only one time window containing a significant effect was observed but this time window contained what appeared to be different components due to their differing topographies, it was decided to conduct more than one model to account for the topographical differences observed. Therefore, a separate linear regression model was conducted for each of the three anteriority regions: anterior, central and posterior. In all three models, the predictors Education and Foreign-accent exposure were excluded, as they were correlated with other predictors (Education was associated with Age [a t-test comparing the ages of participants in

[5] While the ANOVA for the central electrodes showed no significant interactions of Grammaticality by Speaker or Presence of *-s*, examination of the topography plots in Figure 4.1 suggests that the effect observed in response to the native speaker's commission errors at the anterior and posterior electrodes extends across the central electrodes as well, rather than existing as two topographically isolated effects.

the ‘HSC only[6]’ education group (n = 19) and the ‘higher than HSC’ education group (n = 11) revealed that the participants who had completed a higher level of education than the HSC were significantly older that those who reported the HSC as their highest level of education ($t[12.78] = -6.76$, $p < 0.001$)] and Foreign-accent exposure was correlated with both Age [$R^2 = 0.18$, $p = 0.017$] and Mandarin-accent exposure [$R^2 = 0.36$, $p < 0.001$]).

For the model using mean amplitude difference across the anterior electrodes as the dependent variable, Violation type was found to be a significant predictor ($R^2 = 0.03$, $p = 0.027$; commission errors were coded as 1, omission errors as -1). Violation type was also a significant predictor when modelling mean negativity amplitude across the central electrodes ($R^2 = 0.03$, $p = 0.021$). Plotting the data for the two electrode sites revealed that the presence of a commission error predicted a more negative response at the anterior electrodes than the presence of an omission error (Figure 4.2), while omission errors predicted a more negative response at the central electrodes than commission errors (Figure 4.3).

Mean negativity amplitude across the posterior electrodes was not significantly predicted by any of the factors, however the effect of Empathy quotient was marginally significant ($R^2 = 0.02$, $p = 0.057$). A post hoc correlation revealed a trend towards a negative relationship between Empathy quotient and component amplitude at the posterior electrodes ($R^2 = -0.17$). As EQ scores increased (i.e. the participant showed greater empathy), amplitude of the component observed at the posterior electrodes became more negative (Figure 4.4).

In sum, the investigation of individual differences revealed only one between-subjects factor that predicted component amplitude in this study, and this was only a trend, not a significant effect. Participants’ empathy levels were found to modulate their neural responses to morphosyntactic violations at the posterior electrodes. Aside from this, the within-subjects factor of violation type predicted component amplitudes at the anterior and central electrodes, though this is not strictly an example of individual variability as it is a within-subjects factor.

[6] The Higher School Certificate, or HSC, represents the completion of secondary schooling in New South Wales. Participants’ education levels were divided into ‘HSC only’ and ‘higher than HSC’ groups, rather than being treated as a continuous variable, due to the disproportionate number of participants reporting ‘HSC’ as their highest education level (n = 19). Participants in the ‘higher than HSC’ group reported their highest education level as ‘diploma’ (n = 2), ‘associate degree’ (n = 1), ‘Bachelor’s degree’ (n = 6), ‘Honours degree’ (n = 1) or ‘Master’s degree’ (n = 1).

Figure 4.2 – Graph showing relationship between component amplitude and violation type at the anterior electrodes.

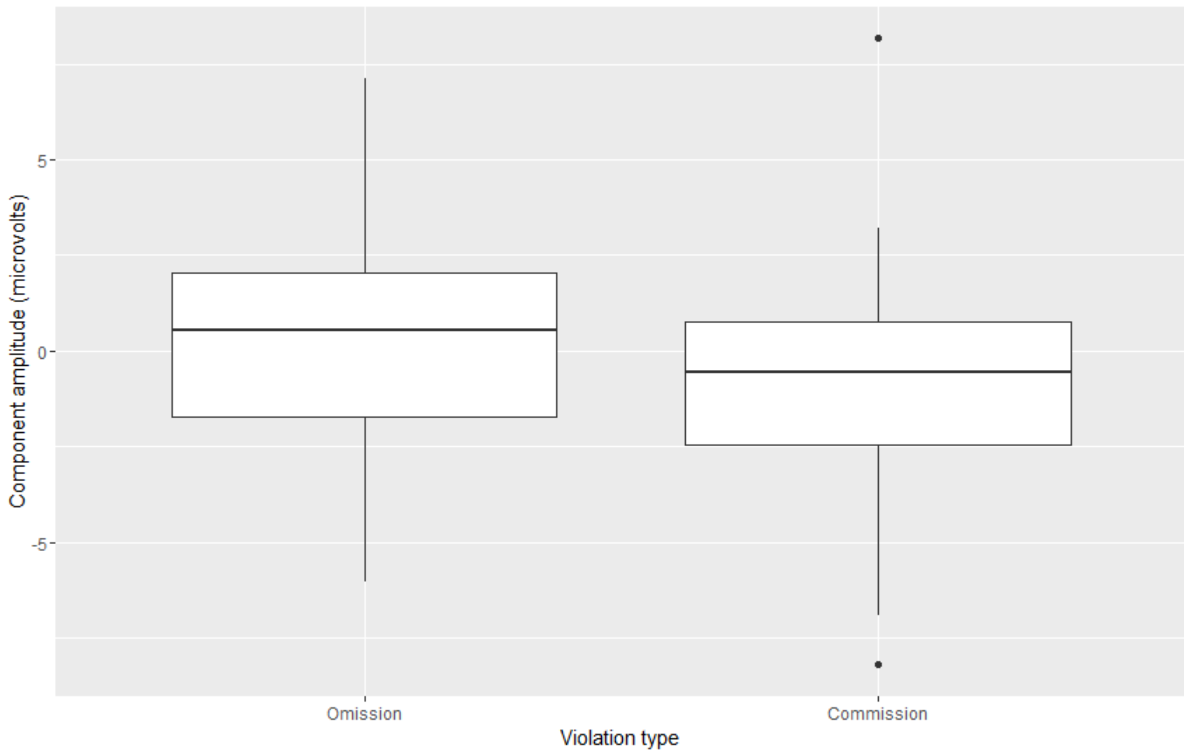


Figure 4.3 – Graph showing relationship between component amplitude and violation type at the central electrodes.

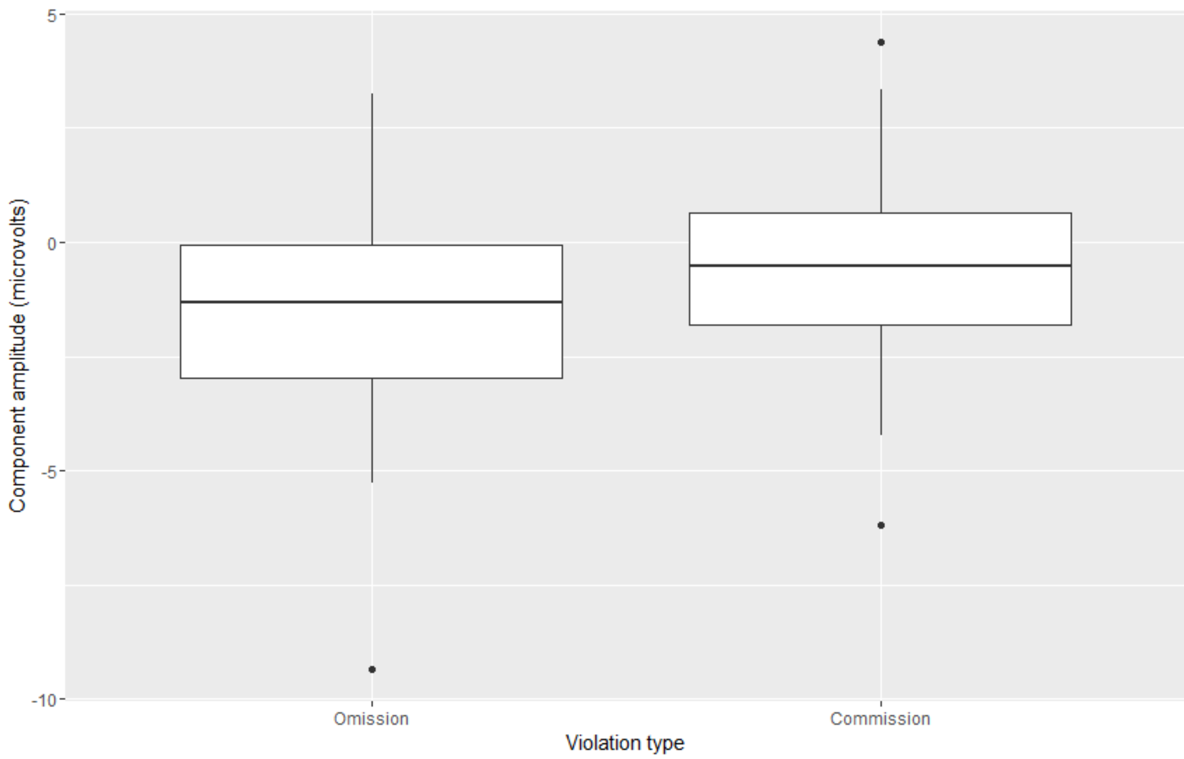
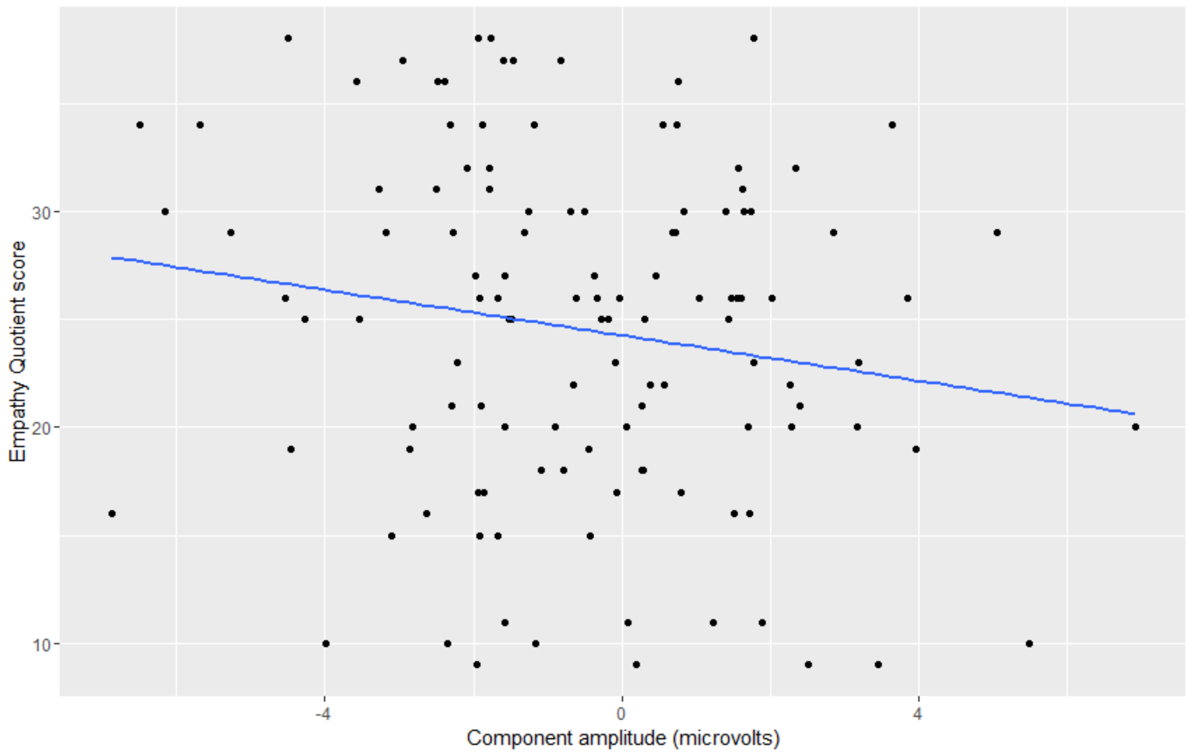


Figure 4.4 – Graph showing relationship between component amplitude and empathy quotient scores at the posterior electrodes.



4.3. Semantic violations

Sentences including semantic violations and their semantically correct counterparts were included in the study in order to demonstrate that the speech of both the native and foreign-accented speakers was being understood and processed by the listeners. It was expected that neural responses to the semantic violations should be more negative than for the corresponding correct sentences, representing a typical N400 effect for semantic violations, and that the presence of this component should not vary across speakers. If an N400 was observed in response to both speakers, this would be taken as evidence that listeners were able to understand and process the speech of both.

Figure 4.5 – Waveform and topography plots for semantic-condition responses.

Waveforms: Native-speaker responses shown in blue, foreign-accented speaker responses in red. Semantically-correct responses shown as solid lines, semantic-violation responses as broken lines. Negativity plotted downwards. Time windows for analysis shown as grey boxes.

Topography plots: Electrodes contributing to the significant clusters are shown as filled black circles.

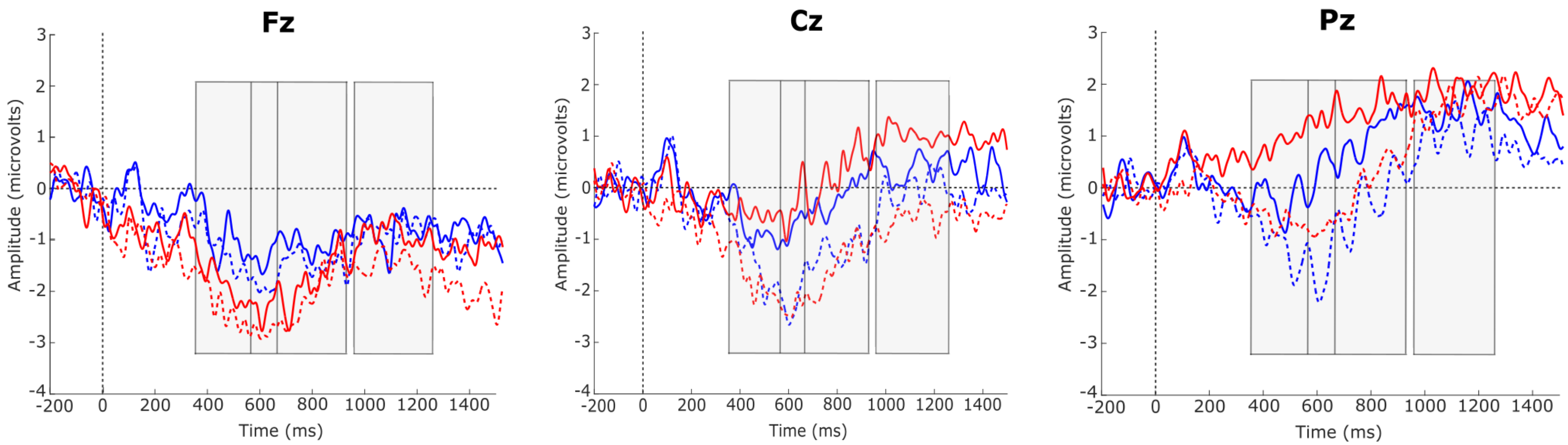
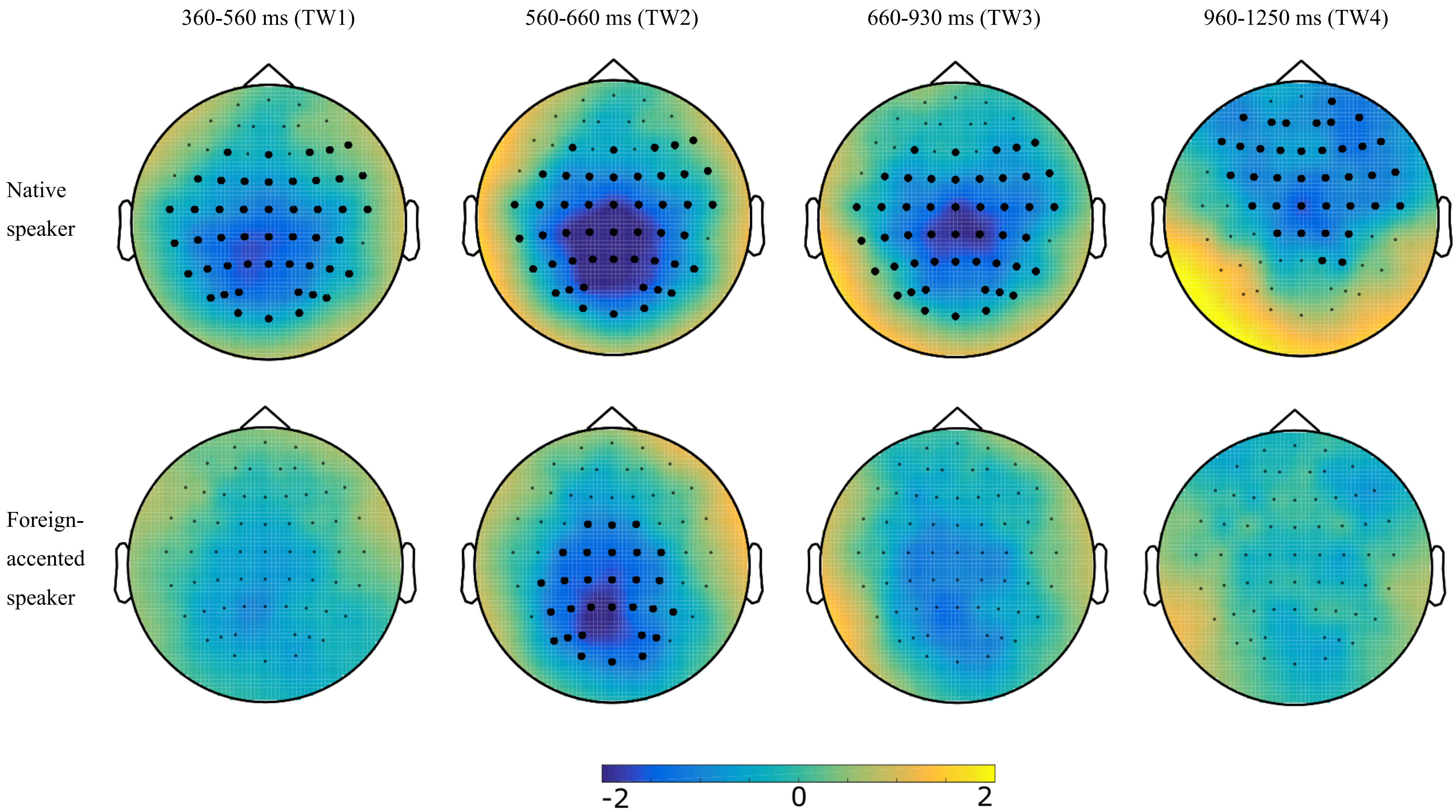


Figure 4.5 cont'd



Visual inspection of the grand average waveforms observed in response to the semantic-condition sentences suggested that a more negative response was elicited by the sentences containing semantic violations than the semantically correct sentences for both speakers in the period approximately 400 to 900 ms after the time-locking point (Figure 4.5). Cluster-based permutation tests confirmed this by revealing a significant difference between the correct and incorrect conditions in response to both speakers. For responses to the native speaker, this difference was reflected by two statistically significant negative clusters. The first lasted from 364 to 931 ms post-stimulus-onset (i.e. post object-noun onset) and was distributed predominantly over the central and posterior electrodes ($p = 0.001$), while the second lasted from 955 to 1245 ms post-stimulus-onset and was distributed across the anterior and central electrodes ($p = 0.012$). In contrast, the difference between the semantically correct and incorrect conditions for the foreign-accented speaker was reflected by one significant negative cluster lasting from 564 to 657 ms post-stimulus-onset, distributed across the central and posterior electrodes ($p = 0.035$).

As the significant clusters in the two speaker conditions displayed very different latencies and durations, four separate time windows were defined to enable comparison between the two speaker conditions. These time windows were as follows: 360 to 560 ms (the period in which the first cluster in response to the native speaker had reached significance but the response to the foreign-accented speaker had not), 560 to 660 ms (the overlap period in which significant clusters were observed in response to both speakers), 660 to 930 ms (the period in which the first cluster in response to the native speaker remained significant but the response to the foreign-accented speaker did not) and 960 to 1250 ms (the second cluster in response to the native speaker). As multiple time windows had now been defined, Time window was added as a factor to the planned ANOVA design, resulting in a 2x2x3x3x4 design: Correctness (correct or incorrect) by Speaker (native or foreign-accented) by Anteriority (anterior, central or posterior) by Laterality (left, midline or right) by Time window (360-560 ms [TW1], 560-660 ms [TW2], 660-930 ms [TW3] or 960-1250 ms [TW4]). As we were only concerned with demonstrating the presence of a significant difference between responses to the correct and incorrect sentences and showing that this effect was not modulated by Speaker, only the key factors of Correctness and Speaker were investigated in detail. We therefore report only the main effects and interactions that involved at least one of these two factors.

The omnibus ANOVA conducted on the semantic-condition data yielded the following results concerning the key variables of Correctness and Speaker (Table 4.1). A significant main effect of Correctness only was observed. Two-way interactions of Correctness by Laterality, Anteriority and Time window were observed, and a two-way interaction of Speaker by Anteriority. No interaction of Correctness by Speaker was observed. Further three- and four-way interactions were also observed: there were interactions of Correctness by Laterality by Anteriority, Correctness by Laterality by Time window, Correctness by Anteriority by Time window, Speaker by Laterality by Time window and Correctness by Laterality by Anteriority by Time window.

Table 4.1 – Significant results involving the key factors Correctness and Speaker shown by the omnibus ANOVA on semantic-condition data.

	(df_1, df_2)	F	p	η^2
Correctness	(1, 29)	11.44	0.002	0.018
Correctness x Laterality	(2, 28)	6.19	0.006	0.006
Correctness x Anteriority	(2, 28)	4.30	0.024	0.002
Correctness x Time window	(3, 27)	6.43	0.002	0.003
Speaker x Anteriority	(2, 28)	9.81	0.001	0.018
Correctness x Laterality x Anteriority	(4, 26)	3.90	0.013	0.001
Correctness x Laterality x Time window	(6, 24)	2.75	0.035	0.001
Correctness x Anteriority x Time window	(6, 24)	4.77	0.002	0.004
Speaker x Laterality x Time window	(6, 24)	3.79	0.009	< 0.001
Correctness x Laterality x Anteriority x Time window	(12, 18)	3.32	0.011	< 0.001

In order to better understand the significant interactions occurring in the semantic-condition data, a separate ANOVA for each level of Time window was conducted. The results of these, again focusing only on main effects and interactions involving Correctness and Speaker, were as follows (Table 4.2). A significant main effect of Correctness was observed in the first three time windows, but not the fourth. The same three two-way interactions were observed in time windows one, two and three: Correctness by Anteriority, Correctness by Laterality and Speaker by Anteriority. No two-way interactions were significant in time window four. A three-way interaction of Correctness by Anteriority by Laterality was observed across time windows two, three and four, and an additional three-way interaction of Correctness by Speaker by Anteriority was observed in time window four.

Table 4.2 – Significant results involving the key factors Correctness and Speaker shown by the individual time-window ANOVAs on semantic-condition data.

	(df_1, df_2)	Time window 1 (360-560 ms)			Time window 2 (560-660 ms)			Time window 3 (660-930 ms)			Time window 4 (960-1250 ms)		
		F	p	η^2	F	p	η^2	F	p	η^2	F	p	η^2
Correctness	(1, 29)	74.45	0.014	0.022	104.24	< 0.001	0.046	22.60	0.002	0.024	-	-	-
Correctness x Anteriority	(2, 28)	6.62	0.004	0.008	9.69	0.001	0.012	4.28	0.002	0.002	-	-	-
Correctness x Laterality	(2, 28)	5.07	0.013	0.007	11.40	< 0.001	0.013	4.60	0.019	0.006	-	-	-
Speaker x Anteriority	(2, 28)	14.95	< 0.001	0.047	11.07	< 0.001	0.029	4.98	0.014	0.014	-	-	-
Correctness x Anteriority x Laterality	(4, 26)	-	-	-	3.45	0.002	0.001	4.90	0.004	0.001	5.19	0.003	0.002
Correctness x Speaker x Anteriority	(2, 28)	-	-	-	-	-	-	-	-	-	4.93	0.015	0.006

Results of the analysis of semantic-condition data, focusing on effects and interactions of Correctness and Speaker, therefore demonstrated that neural responses to sentences containing semantic violations were significantly more negative than responses to their correct counterparts throughout the first three time windows (i.e. from 360 to 930 ms post-stimulus-onset). There was no significant interaction between Correctness and Speaker in any time window, although a higher-order interaction for effects of Correctness, Speaker and Anteriority was observed in time window four, corresponding temporally to the second negative cluster observed in response to sentences produced by the native speaker. It was determined that this level of statistical analysis provided sufficient evidence that a similar negativity was observed in response to semantic violations produced by both speakers between 360 and 930 ms post-stimulus-onset (see section 5.4 for a discussion). Thus further statistical analyses were not undertaken on the semantic-condition data, as the purpose for including this condition was to demonstrate that such a negativity was present.

5. Discussion

The current study used the ERP technique to investigate neural responses to subject-verb agreement violations in native and Mandarin-accented English across two violation types: omission and commission errors. Three hypotheses were proposed. First, a P600, and possibly a LAN, were expected to occur in response to both omission and commission subject-verb agreement errors produced by the native speaker of Australian English. Second, neural responses to those same errors when produced by the Mandarin-accented speaker of English would differ from those in response to the native speaker: either a P600 and possible LAN would be observed in response to commission errors but not omission errors, or neither a P600 nor a LAN would be observed in response to either error type. Third, there would be individual differences between participants' neural responses, which could potentially be related to variability in participants' empathy levels, level of exposure to foreign-accented or Mandarin-accented English, their working memory capacity, how accented they perceived the foreign-accented speaker's speech to be, or whether they were able to identify the accent heard during the task.

Results showed that between 210 and 330 ms post-stimulus-onset, a negativity with a broad distribution was observed in response to commission errors produced by the native speaker. For the omission errors produced by the same speaker, no effect was observed. For commission errors produced by the foreign-accented speaker, a negativity confined to the anterior electrode region was observed. Regarding individual variability within the observed components, the within-subjects factor of violation type was found to predict component amplitude at the anterior and central electrodes, with opposite effects in these two electrode regions. The between-subjects factor of participant empathy was found to have a weak relationship to the amplitude of the component observed at the posterior electrodes, as the effect approached, but did not reach, significance.

Additionally, neural responses to semantic violations were examined as a control condition. This was intended to confirm that listeners were indeed processing the sentences produced by the foreign-accented speaker and were able to comprehend the accented speech (cf. Hanulikovà et al., 2012). N400s were expected to occur in response to semantic violations produced by both speakers, although potentially with variation in latency and/or topography between speakers. In response to the semantic violations a negativity was indeed observed between 360 and 930 ms post-stimulus-onset for both speakers.

In the following discussion, neural responses to the morphosyntactic violations will first be discussed, addressing the first and second hypotheses of the study by examining responses to the native and foreign-accented speaker respectively. The third hypothesis of the study will then be addressed in the discussion of individual differences in the data. Then responses to the semantic-condition sentences will be discussed to provide additional validation for the interpretation of the morphosyntactic-condition sentences. Finally, implications and limitations of the current study and possible directions for future research will be discussed.

5.1. Morphosyntactic violations: Native speaker

The morphosyntactic-condition sentences were the main focus of this study and allowed us to address our aim of determining how speaker identity information, as conveyed by a foreign accent, may affect morphosyntactic processing. Neural responses to the sentences produced by the native speaker will be discussed first, as they address the first hypothesis of the study.

A significant, broadly-distributed negativity between 210 and 330 ms post-stimulus-onset was observed in response to commission-type subject-verb agreement violations in the native speaker's speech. We interpret this negativity as an N400 effect, motivated by the polarity and topographical distribution: the greatest negativity is observed over the centro-parietal electrodes (Figure 4.1), which matches the typical distribution of the N400 (Swaab et al., 2012). While the time window in which our N400 is observed does not seem to match the typical N400 time window, it is worth noting that, firstly, N400 effects in auditory paradigms may be observed as early as 100 ms after stimulus onset (Swaab et al., 2012), and that the time-locking point in N400-eliciting paradigms is generally located at the onset of the stimulus word. In the current study the time-locking point is located near the end of the word, at the beginning of the burst release of the verb-stem-final stop, potentially contributing to the apparently short latency of the N400 effect observed.

The finding of an N400 effect for commission errors in native speech contradicts the first hypothesis of the study, which proposed that both types of subject-verb agreement error would elicit a P600, and possibly a LAN, when produced by the native speaker. This hypothesis was motivated by existing findings that subject-verb agreement violations in English typically elicit a biphasic LAN-P600 response, although the presence of the LAN had been shown to vary according to subtle experimental variations (Osterhout & Mobley, 1995; Molinaro et al., 2011; Dube et al., 2016).

This is not the first study, however, to report an N400 in response to subject-verb agreement errors. This finding has also been reported by Severens, Jansma and Hartsuiker (2008) for subject-verb agreement errors in Dutch. In Severens and colleagues' study, the stimuli were composed of sentences containing complex noun phrases, such as *The street near the church is beautiful*. Each sentence contained both a head noun (*the street*) and a local noun (*the church*). The effect observed was modulated by error type: an N400 occurred when both head and local nouns were singular and a subject-verb agreement error was created by the presence of a plural verb, but a P600 was observed when the sentence included a singular head noun with both a plural local noun and a plural verb.

Severens and colleagues suggested that the N400 and P600 effects observed for the different error types could be explained by the reparability of the violation. They proposed that in the sentences in which the head nouns were singular and the local nouns were plural, presentation of a morphosyntactically incorrect plural verb would trigger reanalysis in an attempt to create a more felicitous parse of the sentence. This reanalysis process, reflected by the presence of the P600, was enabled as the competing local noun displayed an inflection suggesting it could possibly agree with the verb. In the sentences in which both head and local nouns were singular, however, the presentation of a plural verb was irresolvable. The morphosyntactic violation in this circumstance could not be repaired by further syntactic processing as no possible plural subject was present. In this circumstance, syntactic reanalysis, which would be expected to elicit a P600, was not carried out. An N400 was instead observed as a violation of the expected word had occurred.

Severens and colleagues' explanation could be applied to the N400 effect observed in the current study. The morphosyntactic violations investigated in the current study were embedded in sentences of a simple subject-adverb-verb structure in which only one noun was present in the subject noun phrase (Appendix A). This rendered a more felicitous parse impossible as there was no other potential subject noun available, as in Severens and colleagues' singular head noun/singular local noun sentences. The morphosyntactic violations were therefore irresolvable and participants thus made no attempt to resolve them. This resulted in the lack of a P600 component, and instead an N400, representing the processing of an unexpected word, or potentially the unexpected semantic content of the third person singular *-s* morpheme.

While this explanation accounts well for the findings observed in both Severens and colleagues' (2008) study and the current study, Severens and colleagues did raise two potential objections to the interpretation they gave of their results. Firstly, they noted that a previous study had examined subject-verb agreement violations using sentences of the same structure as those in their study and had found the expected P600 in response to both violation types, rather than only the singular head noun/singular local noun sentences. However, they argued that this was likely due to the use of a grammaticality judgement task during EEG recording in the previous study. This had likely prompted participants to pay extra attention to the syntactic structure of the sentences in order to correctly determine whether a violation was present or not. This finding is in line with other previous studies showing task-based effects on the presence or amplitude of the P600 (e.g. Hahne & Friederici, 2002).

Secondly, Severens and colleagues raised the issue that even in previous studies of subject-verb agreement violations that used sentences with simple noun phrases and no grammaticality judgement task, a P600 rather than an N400 was typically observed. However, these studies typically used stimuli with a variety of syntactic structures, as opposed to Severens and colleagues' study, in which every sentence in the task had the same syntactic structure. For this reason they argued that listeners in their study may have engaged in shallow syntactic processing, as the syntactic structure of all the sentences was identical. In contrast, the variety of syntactic structures in the other studies would have required the listeners to engage in deeper syntactic (re)analysis, as the structure of each sentence was not as predictable.

Crucially, both arguments presented by Severens and colleagues (2008) can be applied to the current study. The current study does not employ a grammaticality judgement task, and thus does not force listeners to engage in syntactic reanalysis. The current study also uses stimuli in which the syntactic structure is kept more-or-less constant, even across the morphosyntactic and semantic conditions (i.e. subject + [adverb] + verb + [object]). While the syntactic structure of the fillers in the current study was not controlled like that of Severens and colleagues' study, the requirement that all fillers be six or seven syllables in length did not allow for a great deal of syntactic variation, further contributing to the structurally repetitive nature of the stimuli. This structural consistency may have been sufficient to induce shallow syntactic processing. Therefore, we attribute the finding of an N400 in response to both omission and commission errors produced by the native speaker in the current study to the irreparability of the violation and the use of shallow syntactic processing. Because no alternate parse was available to participants, and because there was no incentive for them to conduct deep syntactic analysis, participants did not reanalyse the sentences containing morphosyntactic violations but instead treated the violation as semantic, resulting in the presence of an N400 but neither a LAN nor a P600.

It remains somewhat unclear why other studies investigating subject-verb agreement errors have not found an N400 effect for the reasons described above. For example, in the study of Dube and colleagues (2016) it was also the case that stimuli sentences contained only one noun, so no alternate parse was available. Also, no grammaticality judgement task was used, and the syntactic structure of the sentences was kept constant, even more so than in the current study. Assuming the interpretation of the N400 proposed by Severens and colleagues (2008) is sound, why did Dube and colleagues observe a biphasic LAN-P600 and not an N400? One potential explanation is that the inclusion of sentences containing semantic violations in the stimuli of the current study may have biased the participants towards favouring a more semantics-based processing strategy. Participants may have been more willing to treat the presence of an incorrect morpheme or the absence of a correct morpheme as an incongruous word choice, rather than a grammatical violation, due to their exposure to highly salient semantic violations throughout the task.

A further question arising from the results of the native speaker's morphosyntactic violations is the finding of an N400 effect for the commission errors only and not the omission errors. It was expected that neural responses to both types of subject-verb agreement error would be similar for the native speaker, given that listeners would expect the native speaker to be equally good at constructing subject-verb agreement in both singular- and plural-subject contexts. Listeners were thus expected to find violations of both types equally unacceptable and show similar neural responses across error types. However, the lack of effect observed in the omission-error context may best be explained, not with reference to the listener's expectations, but rather the relative perceptual salience of the two violation types.

Dube and colleagues (2016) argue, in their discussion of English omission and commission subject-verb agreement errors, that commission errors are more perceptually salient due to their acoustic characteristics. The high-frequency noise associated with the phoneme /s/ is a salient overt marker of a commission error in the speech stream, whereas an omission error has no overt acoustic correlate (its correlate is rather a lack of acoustic information). Thus it can be argued that commission errors are a more salient error type than omission errors. This may explain why an N400 effect was observed in response to the commission errors but not the omission errors in the current study, as the omission violations' relative lack of salience may be responsible for the lack of a violation effect.

To summarise, listeners displayed an N400 effect in the 210 to 330 ms time window in response to errors of commission produced by a native speaker. However, no such effect was observed for omission errors. This difference can be explained by the relative perceptual salience of the two error types. These findings contradicted the first hypothesis of the study, which proposed that a P600 component and possible LAN would be observed. The finding of an N400 in the absence of a LAN or P600 was most likely due to shallow syntactic parsing driven by the irreparability of the morphosyntactic violations and the lack of incentive provided to the participants to engage in deep syntactic processing. Additionally, the inclusion in the task of sentences containing salient semantic violations may have primed the listeners to process the morphosyntactic violations in a 'semantic' manner.

5.2. Morphosyntactic violations: Foreign-accented speaker

Compared to those for native speech, a different pattern of neural responses was observed for morphosyntactic violations in foreign-accented speech. Commission errors produced by the foreign-accented speaker elicited a negativity across the anterior electrodes in the 210-330 ms time window, with no effect observed for omission errors.

The effect observed in response to commission errors produced by the foreign-accented speaker, while occurring at the same latency as the N400s described in section 5.2, was distributed across the anterior electrodes only (Figure 4.1). This distribution distinguishes it from the centro-parietally maximal N400 observed in response to the native speaker's commission errors (Figure 4.1). We interpret this negativity as a working-memory-related AN. As discussed in section 2.2.3, the LAN/AN may be interpreted either as a response to morphosyntactic violations, or as an index of increased working memory load. The topography of the observed AN seems to most closely match the topography found in studies where the working memory load involved in processing the stimuli is great (e.g. Piai, Meyer, Schreuder, & Bastiaansen, 2013). This, together with the fact that the AN does not display the left-lateralisation associated with the canonical LAN elicited by morphosyntactic violations, motivates our working-memory interpretation.

It therefore seems that commission errors produced by the foreign-accented speaker do not elicit a violation effect. However, listeners do make extra use of their working memory in this circumstance, potentially revisiting earlier parts of the sentence to determine whether or not a violation did actually occur. This is very different to the violation response observed for the same errors produced by the native speaker. It seems that when hearing a commission error produced by a native speaker, the listener immediately responds to it as a violation, however when the commission error is produced by a foreign-accented speaker, the listener must engage their working memory to determine whether a violation is present or not.

The effect observed in response to the foreign-accented speaker's omission errors, however, matches its native-speaker counterpart: no effect is observed. As with the native-speaker results discussed above, we attribute the lack of response to omission errors to their relative lack of perceptual salience compared to the commission errors.

The neural responses observed for the morphosyntactic violations produced by the foreign-accented speaker differ in several respects from what was hypothesised. The second hypothesis of the study proposed that listeners would show different neural responses to subject-verb agreement errors in foreign-accented speech compared to their counterparts in native speech. However, the precise details of the hypothesis were contingent upon the results observed for the native speaker. Supposedly, both omission and commission errors in native speech would have elicited a P600 and/or LAN. If this was indeed the case, two possible scenarios could occur for the errors in the foreign-accented speech. First, only commission errors, but not omission errors, would elicit a P600 and/or LAN. In this case the difference in responses to the two speakers could be taken as evidence that participants' neural responses were indeed modulated by speaker-identity information based on participants' prior experience with Mandarin-accented English. Alternatively, it could be the case that neither violation type would yield a P600 and/or LAN. This would again point to an effect of speaker identity on morphosyntactic processing, but independent of prior experience. However these two possible scenarios relied on the assumption that a P600 and/or LAN would be observed for the native speaker in the first place.

Considering that an N400, rather than a LAN or P600, was observed in response to both error types for the native speaker, the details of the two proposed scenarios must be updated to allow for these findings. Therefore, we will take any significant difference between neural responses to violations made by the two speakers as evidence for speaker-identity integration in morphosyntactic processing. If this difference is observed for the omission errors but not commission errors, this will provide support for the first scenario: that the difference is driven by participants' expectations based on prior experience with Mandarin-accented speech. If the difference is observed for both error types, this will support the second scenario: that the effect is instead a general effect of foreign-accentedness and is not based on experience.

The neural responses observed for the native and foreign-accented speakers were indeed shown to differ, demonstrating that speaker-identity integration does occur during the processing of subject-verb agreement errors in English. Listeners displayed an N400 effect when presented with commission errors produced by a native speaker of Australian English and an AN when presented with commission errors produced by a Mandarin-accented speaker of English. Because the stimuli used in these two conditions did not differ apart from the speaker-identity information conveyed by the speaker's accent, we can conclude that speaker identity, as conveyed by a foreign accent, is modulating the neural responses observed.

However, when it comes to the two proposed scenarios for the manner in which responses to the two speakers may differ, the interpretation of the results becomes less clear. It was not the case that neural responses differed between speakers for the omission but not the commission violations, which would suggest an effect of prior experience. Neither was it true that neural responses differed in response to both violation types, which would suggest otherwise. Instead, neural responses differed for the commission violations, but remained the same across speakers for the omission violations. Therefore, despite having demonstrated that speaker identity does influence the neural correlates of subject-verb agreement processing, we are unable to make a claim either for or against the proposition put forward by Hanulikovà and colleagues (2012), suggesting that the effect of a foreign accent on morphosyntactic processing is driven by expectations stemming from prior experience with that accent.

It could potentially be argued that the difference in ERP patterns observed for the commission errors produced by the native and foreign-accented speakers is due to the lower intelligibility of the foreign-accented speech, rather than the integration of speaker-identity information. This is particularly the case since the processing of foreign-accented speech has previously been found to place greater demands on working memory than processing native speech (Van Engen & Peelle, 2014) and the component in question is associated with working memory use. However, no working-memory component (i.e. an AN) was observed in response to either the omission errors or semantic errors (see section 5.4) produced by the foreign-accented speaker. If this working-memory effect was a general response to processing difficulty associated with foreign-accented speech, we would expect it to occur in all circumstances in which the participants are exposed to foreign-accented speech. As this is not the case, it seems reasonable to instead interpret the presence of this effect as being related to the integration of speaker-identity information.

The question remains as to why it should be that speaker-identity integration seems to only occur in the processing of commission rather than omission errors. There are two possible reasons why this could be the case. Firstly, speaker-identity integration may be more likely to occur in response to more salient error types. Listeners may only bother integrating the speaker's identity when a highly salient error must be processed. Alternatively, a more robust violation effect might be necessary in order to observe the effect of speaker-identity integration on neural responses. That is, speaker-identity integration may be occurring here, even for the omission errors, but the strength of the violation effect for both speakers is not sufficient for any differences between the two to become apparent.

Thus the neural responses observed to subject-verb agreement violations in foreign-accented speech provide support for the findings of Hanulikovà and colleagues (2012) in that an effect of foreign-accentedness is observed, at least for commission errors. This demonstrates that speaker-identity integration does indeed occur in English morphosyntactic processing, and for subject-verb agreement errors (at least of the commission type). However, the role that prior experience plays in speaker integration remains unclear from the current study, given the lack of any significant difference between neural responses to the two speakers when the errors were of the omission type.

5.3. Individual differences

Aside from the group level effects described in sections 5.1 and 5.2, a small amount of individual variability was also observed in the data, supporting the third hypothesis of the study. ERP effect amplitudes at various regions of the scalp were found to be predicted by different factors.

Firstly, violation type was found to predict the component amplitude observed at the anterior and central electrodes. Commission errors were shown to predict greater negativity amplitude at the anterior electrodes, while omission errors were shown to predict greater negativity amplitude at the central electrodes. This observation, while useful in corroborating the results of the analyses presented in sections 5.1 and 5.2, is not particularly helpful in the discussion of individual differences in responses between participants, as the type of violation was a within-subjects, rather than between-subjects, factor.

The only between-subjects factor that was shown to have even a marginally significant effect on the amplitude of the components observed was participant empathy (indexed by Empathy Quotient score). As participants' empathy increased, their responses to morphosyntactic violations at the posterior electrodes followed a trend towards becoming more negative. This initially seems counterintuitive, as we might expect the more empathic participants to be more accepting of errors, particularly those produced by the foreign-accented speaker. However, commission errors produced by the native speaker were the only error type that elicited a significant effect at the posterior electrodes. This suggests that participants' empathy influences their neural response to the native speaker's commission errors, rather than any response to the foreign-accented speaker. However, the N400 effect for native-speaker commission errors was observed across all three anteriority regions – why should it only be predicted by participants' empathy at the posterior electrodes? This can be explained by the overlap of the N400 and AN at the more anterior electrode sites. The overlapping effects may have obscured the potential relationship between N400 amplitude and empathy.

While this finding supports the broad hypothesis that there may be an effect of participant empathy on neural responses, it contrasts with the more detailed hypothesis presented in Table 2.1. There it was suggested that greater empathy would result in a smaller amplitude for the violation component observed in response to the foreign-accented speaker's errors. This hypothesis was not borne out by the results, as the amplitude modulation was actually observed for responses to the native speaker rather than the foreign-accented speaker. The directionality of the effect was also reversed, as greater empathy was shown to increase, rather than decrease, component amplitude.

Why should more empathic participants display a stronger N400 in response to the native speaker's commission errors? A potential reason is that more empathic participants, all native speakers of Australian English themselves, are aware that subject-verb agreement errors, and highly salient commission errors in particular, should not be present in the speech of a native Australian English speaker. They are able to make strong predictions regarding the cognitive processes of the native speaker, who is similar to themselves, due to their higher levels of cognitive empathy (recalling that the Empathy Quotient score is a measure of both cognitive and affective empathy). This may result in a stronger violation effect when those expectations are violated by the presence of a commission error, causing the observed increase in N400 amplitude. Notably, a similar relationship was observed by van den Brink and colleagues (2012), who found that more empathic participants showed a greater violation effect when processing sentences containing sex-based speaker incongruities.

This relationship between N400 amplitude and participant empathy suggests that, out of all the between-subjects factors included in the linear regression models, empathy may be involved in predicting how individual participants will respond to linguistic stimuli, even though the relationship is not strong. As previously stated, this finding was only marginally significant and thus should best be interpreted as a trend. However, the observation of a potential empathy effect on morphosyntactic processing is a novel finding, as an effect of empathy has previously only been observed in the processing of semantic violations (van den Brink et al. 2012).

5.4. Semantic violations

The semantic-violation condition was included in this study to confirm that participants were understanding and processing the foreign-accented speech presented in the task. If the listeners were able to understand the foreign-accented speech as well as the native speech, this would be reflected in an N400 effect in response to semantic violations made by both speakers. This was supported by the present results.

Results of the cluster-based permutation tests showed a significant difference between the semantically-correct and -incorrect sentences for both speakers. For the native speech, this was in the form of two significant negative clusters: a long-lasting central and posterior negativity lasting from 364 to 931 ms, and a shorter, more frontally-distributed negativity between 955 and 1245 ms post-stimulus onset. For the foreign-accented speaker only one cluster was identified: a central and posteriorly distributed negativity between 564 and 657 ms post-stimulus-onset. However, further statistical analysis using ANOVAs revealed that the significant difference between responses to semantically correct and incorrect sentences lasted from 360 to 930 ms post-stimulus-onset for both speakers. During this time period there was no significant interaction between Correctness and Speaker, or between Correctness, Speaker and either Anteriority or Laterality. This suggests that the negativity observed for foreign-accented speech did not differ from that for native speech in terms of latency, duration, amplitude or scalp distribution.

We interpret these negativities as N400 effects, reflecting difficulties in semantic memory retrieval and/or integration. The observation for similar N400s for both speakers provides evidence that listeners were equally able to process and comprehend the speech of both. This demonstrates that any lack of significant ERP results in other conditions cannot be attributed to difficulty in parsing foreign-accented speech.

5.5. Implications

The results of the current study have two implications for our understanding of sentence processing. Firstly, the current study contributes to the evidence suggesting that the integration of speaker-identity information is integral to sentence processing. While there is already quite some evidence for this from studies in phonology and semantics (see section 2.3), few studies have investigated speaker-identity integration with respect to syntactic/morphosyntactic processing. As the current study has demonstrated that speaker-identity integration does indeed seem to occur in morphosyntactic processing, it contributes to the body of evidence that suggests speaker-identity integration is relevant at every step of sentence processing. This rapid on-line integration of multiple sources of information seems to provide some support for more interactive models of sentence processing (e.g. Pickering & Garrod, 2013).

Secondly, the findings of this study suggest that the relationship between participants' empathy levels and their neural responses to linguistic stimuli may occur in broader contexts than what have so far been described. The relationship between empathy and linguistic processing has been examined in with respect to semantic processing previously (van den Brink et al. 2012), however this study provides a suggestion that it may also play a role in predicting responses to morphosyntactic violations. However, further research will be required to verify this claim, as the relationship observed in this study was weak and did not reach significance, and participants may have been using a 'semantic' processing strategy for the morphosyntactic violations.

5.6. Limitations

While this study has been able to demonstrate that the integration of speaker identity information, as conveyed by a foreign accent, does impact the processing of subject-verb agreement errors in English, it does suffer from some limitations.

Firstly, the results of the study may have been adversely affected by the stimuli design. The simplicity and structural similarity of the sentences may have facilitated listeners' use of shallow syntactic processing. This may have potentially obscured effects which would have become significant had deeper syntactic (re)analysis taken place. Similarly, the presence of semantic violations in the stimuli may have influenced the listeners' processing strategies, thus potentially skewing the data towards a more 'semantic' treatment of the morphosyntactic violations.

Secondly, the location of the baseline period in the morphosyntactic condition was far from the trigger point, which could be argued to cause problems in the data. However this choice of baseline period was data-driven, rather than being an arbitrary selection: cluster-based permutation tests revealed that the period chosen displayed no significant differences between conditions, whereas other potential baseline locations did (see Appendix E). Thus in the current study this potential limitation has been mitigated as far as possible.

Thirdly, the linear regression models used to examine potential sources of individual variation in the data also suffer from some limitations. No interaction terms were included in these models to assist in interpretability due to the large number of factors present in each model, and no random factors were included. This may have limited the predictive power of the models by limiting the data available. Also, a more informative and coherent picture of the data may have emerged if component amplitudes from all three anteriority regions had been entered into a single model, rather than three separate models, as was the case in the current study. A more comprehensive model that includes random factors of Subject and Item, interaction terms and multiple dependent variables could potentially be used in future. This would likely improve the accuracy and coherence of the results of the investigation of individual variation, and may reveal that additional predictors did have significant effects on component amplitude when interaction terms and all three anteriority regions were taken into account.

5.7. Future directions

There are several avenues for potential future research suggested by this study. While we have demonstrated that speaker-identity integration does have some effect on the processing of subject-verb agreement errors in English, we have not been able to provide evidence regarding whether or not this effect is driven by participants' expectations based on prior experience, as proposed by Hanulikovà and colleagues (2012), or whether this is a general response to foreign-accented speech, independent of real-life experience. Further research will be required to elucidate which of these hypotheses is likely to be correct.

To do this, future studies could utilise aspects of study design which may elicit stronger ERP responses to morphosyntactic violations and prevent participants from using a shallow syntactic processing strategy, as was likely the case in the current study. Stimulus sentences including more than one potential subject noun, as in Severens and collegues' (2008) study, could be used to facilitate reanalysis when a violation is present, and/or stimuli with a greater amount of structural variety could be used to encourage deeper syntactic processing. This may elicit more robust neural responses in all conditions, which could provide more informative results than the current study. Alternatively, a different morphosyntactic construction in which violation types still differ in terms of expectedness but do not differ in terms of perceptual salience could be investigated. A study of this design would be more likely to elicit consistent responses across error types produced by the native speaker, making the impact of speaker-identity information in the foreign-accented speech clearer.

The observation that the integration of speaker-identity information impacts morphosyntactic processing raises the question of how listeners learn to use this speaker-identity information. Future research could therefore also address speaker-identity integration from an acquisition perspective. Researchers have begun examining children's acquisition of speaker-identity integration in semantic processing (Borovsky & Creel, 2014), however further research in this area is required, and extension to morphosyntactic processing has not been explored yet. This avenue could potentially be investigated by utilising an 'alien learning paradigm' (Labelle & Valois, 2003; Courteau et al., 2013).

The observation of a relationship, albeit a small one, between participant empathy and neural responses to subject-verb agreement errors in some circumstances raises the possibility that responses to morphosyntactic violations more generally may also be influenced by listeners' empathy. While the effect of listeners' empathy on speaker-identity integration in semantic processing has been examined previously (van den Brink et al., 2012), it has not been examined in morphosyntactic processing, either with or without the manipulation of speaker identity. Listener's empathy has also not been investigated in relation to purely linguistic semantic violations (i.e. violations caused by production of a completely incongruent word, rather than a word rendered incongruent by the speaker's identity). The results observed in the current study suggest that investigating the relationship between participants' empathy and both pure semantic violations and morphosyntactic violations may be a fruitful area for future study.

Furthermore, this study suggests that the characteristics of all the stimuli presented in an experimental setting can have an effect on the neural responses observed, including sentences from supposedly separate conditions and even filler sentences. By suggesting that the inclusion of semantic violations may have affected how the morphosyntactic violations in the study were processed, the results of the current study highlight the need to account for all the sentences that participants hear, not just the target sentences. Additionally, our results also highlight the effect that stimulus complexity can have on neural responses. While it is necessary to avoid overly complex stimuli, the results of the current study suggest that stimulus variability may be a desirable trait, given that repetitive stimuli seem conducive to shallow processing. Future research into the effects that differing degrees of stimulus complexity or the inclusion of fillers of different types can have on listeners' neural responses would be of great benefit in elucidating the degree to which task effects may be influencing the results of sentence-processing ERP studies.

Finally, the results of the semantic-violation condition also displayed some interesting properties which may warrant further investigation. For example, the results fail to replicate previous findings suggesting that semantic violations produced by a foreign-accented speaker elicit N400s with a smaller amplitude (Goslin, Duffy, & Floccia, 2012) or different topography (Hanulikovà et al., 2012) to those in native speech. Additionally, a second negative phase separated from the previous phase by 30ms has not been reported for semantic violations before. While it is possible that the separation between these two effects could potentially be due to noise in the data, this seems unlikely given the difference in topography between the first and second phases. Further research could address the identity and functional significance of this second negative component.

6. Conclusion

In summary, this ERP study examined neural responses to subject-verb agreement errors in both native and Mandarin-accented English. The purpose of this was to demonstrate whether or not the integration of speaker-identity information, as represented by a foreign accent, affected morphosyntactic processing in English, as it had only previously been shown in Dutch, and whether the effect would be observed for subject-verb agreement errors, as it had only previously been observed for grammatical gender agreement errors. Additionally, examining the combination of subject-verb agreement errors and Mandarin-accented English was intended to elucidate the role of prior real-life experience in speaker-identity integration during morphosyntactic processing.

The results for omission subject-verb agreement errors demonstrated no difference in neural responses to the two speakers. Due to this, it could not be concluded whether or not the patterns of speaker integration in morphosyntactic processing were dependent on prior experience or not. However, it was found that the speaker's accentedness did indeed modulate neural responses to subject-verb agreement errors when the error was of the commission type, demonstrating that speaker integration does have an effect on morphosyntactic processing in English, and does occur with respect to subject-verb agreement. Additionally, a small effect of participant empathy was observed on component amplitude, suggesting that this may be a possible factor influencing individual variation among the neural responses of the participants to morphosyntactic violations.

References

Adank, P., Davis, M. H., & Hagoort, P. (2012). Neural dissociation in processing noise and accent in spoken language comprehension. *Neuropsychologia, 50*(1), 77-84. doi: 10.1016/j.neuropsychologia.2011.10.024

Adank, P., Noordzij, M. L., & Hagoort, P. (2012). The role of planum temporale in processing accent variation in spoken language comprehension. *Human Brain Mapping, 33*(2), 360-372. doi: 10.1002/hbm.21218

Boersma, P., & Weenink, D. (2015). Praat: Doing phonetics by computer (Version 5.4.08).

Bornkessel-Schlesewsky, I., Krauspenhaar, S., & Schlesewsky, M. (2013). Yes, you can? A speaker's potency to act upon his words orchestrates early neural responses to message-level meaning. *PLoS ONE, 8*(7), e69173. doi: 10.1371/journal.pone.0069173

Bornkessel-Schlesewsky, I., Kretzschmar, F., Tune, S., Wang, L., Genç, S., Philipp, M., . . . Schlesewsky, M. (2011). Think globally: Cross-linguistic variation in electrophysiological activity during sentence comprehension. *Brain and Language, 117*(3), 133-152. doi: 10.1016/j.bandl.2010.09.010

Borovsky, A., & Creel, S. C. (2014). Children and adults integrate talker and verb information in online processing. *Developmental Psychology, 50*(5), 1600-1613. doi: 10.1037/a0035591

Chen, H., Xu Rattanasone, N., Cox, F., & Demuth, K. (in submission). Effect of early dialectal exposure on adult perception of phonemic vowel length.

Cohen, M. X. (2014). *Analyzing Neural Time Series Data: Theory and Practice*. Cambridge, MA: MIT Press.

Coulson, S., King, J. W., & Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes, 13*(1), 21-58. doi: 10.1080/016909698386582

Courteau, E., Royle, P., Gascon, A., Marquis, A., Drury, J. E., & Steinhauer, K. (2013). Gender concord and semantic processing in French children: An auditory ERP study. In S. Baiz, N. Goldman, & R. Hawkes (Eds.), *BUCLD 37 Proceedings* (Vol. 1, pp. 87-99). Sommerville, MA: Cascadilla.

Crystal, D. (2003). *The Cambridge Encyclopedia of the English Language* (2nd ed.). Cambridge, UK: Cambridge University Press.

Donchin, E. (1981). Surprise!… surprise? *Psychophysiology, 18*(5), 493-513.

Dube, S., Kung, C., Peter, V., Brock, J., & Demuth, K. (2016). Effects of type of agreement violation and utterance position on the auditory processing of subject-verb agreement: An ERP study. *Frontiers in Psychology, 7*.

Fiebach, C. J., Schlesewsky, M., & Friederici, A. D. (2002). Separating syntactic memory costs and syntactic integration costs during parsing: the processing of German WH-questions. *Journal of Memory and Language, 47*(2), 250-272. doi: 10.1016/S0749-596X(02)00004-9

Floccia, C., Butler, J., Goslin, J., & Ellis, L. (2009). Regional and foreign accent processing in English: Can listeners adapt? *Journal Of Psycholinguistic Research, 38*(4), 379-412. doi: 10.1007/s10936-008-9097-8

Foucart, A., Garcia, X., Ayguasanosa, M., Thierry, G., Martin, C., & Costa, A. (2015). Does the speaker matter? Online processing of semantic and pragmatic information in L2 speech comprehension. *Neuropsychologia, 75*, 291-303. doi: 10.1016/j.neuropsychologia.2015.06.027

Frenzel, S., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2011). Conflicts in language processing: A new perspective on the N400–P600 distinction. *Neuropsychologia, 49*(3), 574-579. doi: 10.1016/j.neuropsychologia.2010.12.003

Friederici, A. D. (1995). The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain and Language, 50*(3), 259-281. doi: 10.1006/brln.1995.1048

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences, 6*(2), 78-84. doi: 10.1016/S1364-6613(00)01839-8

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Cognitive Brain Research, 1*(3), 183-192. doi: 10.1016/0926-6410(93)90026-2

Galfano, G., Mazza, V., Angrilli, A., & Umiltà, C. (2004). Electrophysiological correlates of stimulus-driven multiplication facts retrieval. *Neuropsychologia, 42*(10), 1370-1382.

Ganis, G., Kutas, M., & Sereno, M. I. (1996). The search for "common sense": An electrophysiological study of the comprehension of words and pictures in reading. *Journal of Cognitive Neuroscience, 8*(2), 89-106. doi: 10.1162/jocn.1996.8.2.89

Goslin, J., Duffy, H., & Floccia, C. (2012). An ERP investigation of regional and foreign accent processing. *Brain and Language, 122*(2), 92-102. doi: 10.1016/j.bandl.2012.04.017

Gouvea, A. C., Phillips, C., Kazanina, N., & Poeppel, D. (2010). The linguistic processes underlying the P600. *Language and Cognitive Processes, 25*(2), 149-188. doi: 10.1080/01690960902965951

Hagoort, P. (2003). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of Cognitive Neuroscience, 15*(6), 883-899. doi: 10.1162/089892903322370807

Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes, 8*(4), 439-483. doi: 10.1080/01690969308407585

Hagoort, P., & Brown, C. M. (1999). Gender electrified: ERP evidence on the syntactic nature of gender processing. *Journal Of Psycholinguistic Research, 28*(6), 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*(11), 1531-1549. doi: 10.1016/S0028-3932(00)00053-1

Hagoort, P., & Brown, C. M. (2000). ERP effects of listening to speech: Semantic ERP effects. *Neuropsychologia, 38*(11), 1518-1530. doi: 10.1016/S0028-3932(00)00052-X

Hahne, A., & Friederici, A. D. (2002). Differential task effects on semantic and syntactic processes as revealed by ERPs. *Cognitive Brain Research, 13*(3), 339-356.

Hanulíková, A., & Carreiras, M. (2015). Electrophysiology of subject-verb agreement mediated by speakers' gender. *Frontiers in Psychology, 6*. doi: 10.3389/fpsyg.2015.01396

Hanulíková, A., van Alphen, P. M., van Goch, M. M., & Weber, A. (2012). When one person's mistake is another's standard usage: The effect of foreign accent on syntactic processing. *Journal of Cognitive Neuroscience, 24*(4), 878-887. doi: 10.1162/jocn_a_00103

Hasting, A. S., & Kotz, S. A. (2008). Speeding up syntax: On the relative timing and automaticity of local phrase structure and morphosyntactic processing as reflected in event-related brain potentials. *Journal of Cognitive Neuroscience, 20*(7), 1207-1219.

Hoeks, J. C. J., Stowe, L. A., Hendriks, P., & Brouwer, H. (2013). Questions left unanswered: How the brain responds to missing information. *PLoS ONE, 8*(10), e73594.

Holcomb, P. J., & McPherson, W. B. (1994). Event-related brain potentials reflect semantic priming in an object decision task. *Brain and Cognition, 24*(2), 259-276. doi: 10.1006/brcg.1994.1014

Hoshino, N., Dussias, P. E., & Kroll, J. F. (2010). Processing subject–verb agreement in a second language depends on proficiency. *Bilingualism: Language and Cognition, 13*(02), 87-98. doi: 10.1017/S1366728909990034

Jasper, H. H. (1958). The ten twenty electrode system of the international federation. *Electroencephalography and Clinical Neurophysiology, 10*, 371-375.

Jia, G., & Fuse, A. (2007). Acquisition of English grammatical morphology by native Mandarin-speaking children and adolescents: Age-related differences. *Journal of Speech, Language, and Hearing Research, 50*(5), 1280-1299. doi: 10.1044/1092-4388(2007/090)

Johnson, K., Strand, E. A., & D'Imperio, M. (1999). Auditory–visual integration of talker gender in vowel perception. *Journal of Phonetics, 27*(4), 359-384. doi: 10.1006/jpho.1999.0100

Jolsvai, H., Sussman, E., Csuhaj, R., & Csépe, V. (2011). Neurophysiology of Hungarian subject–verb dependencies with varying intervening complexity. *International Journal of Psychophysiology, 82*(3), 207-216.

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes, 15*(2), 159-201. doi: 10.1080/016909600386084

Kappenman, E. S., & Luck, S. J. (2012). ERP components: The ups and downs of brainwave recordings. In S. J. Luck & E. S. Kappenman (Eds.), *The Oxford Handbook of Event-Related Potential Components* (pp. 3-30). New York, NY: Oxford University Press.

King, J. W., & Kutas, M. (1995). Who did what and when? Using word-and clause-level ERPs to monitor working memory usage in reading. *Journal Of Cognitive Neuroscience, 7*(3), 376-395.

Kos, M., Vosse, T. G., Van Den Brink, D., & Hagoort, P. (2010). About edible restaurants: Conflicts between syntax and semantics as revealed by ERPs. *Frontiers in Psychology, 1*, 222.

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research, 1146*, 23-49.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences, 4*(12), 463-470. doi: 10.1016/S1364-6613(00)01560-6

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62*(1), 621-647. doi: 10.1146/annurev.psych.093008.131123

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*(4427), 203-205.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307*(5947), 161-163.

Kutas, M., Neville, H. J., & Holcomb, P. J. (1987). A preliminary comparison of the N400 response to semantic anomalies during reading, listening and signing. *Electroencephalography and Clinical Neurophysiology Supplement, 39*, 325-330.

Labelle, M., & Valois, D. (2003). Floated quantifiers, quantifiers at a distance, and logical form constructions in the acquisition of L1 French. In B. Beachley, A. Brown, & F. Conlin (Eds.), *BUCLD 27 Proceedings* (Vol. 2, pp. 473-483). Somerville, MA: Cascadilla.

Ladefoged, P., & Broadbent, D. E. (1957). Information Conveyed by Vowels. *The Journal of the Acoustical Society of America, 29*(1), 98-104. doi: 10.1121/1.1908694

Lattner, S., & Friederici, A. D. (2003). Talker's voice and gender stereotype in human auditory sentence processing: Evidence from event-related brain potentials. *Neuroscience Letters, 339*(3), 191-194. doi: 10.1016/S0304-3940(03)00027-2

Luck, S. J. (2005). *An Introduction to the Event-Related Potential Technique*. Cambridge, MA: MIT Press.

Major, R. C. (2001). *Foreign accent: The ontogeny and phylogeny of second language phonology*. Mahwah, NJ/London, UK: Lawrence Erlbaum Associates.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of Neuroscience Methods, 164*(1), 177-190.

Molinaro, N., Barber, H. A., & Carreiras, M. (2011). Grammatical agreement processing in reading: ERP findings and future directions. *Cortex, 47*(8), 908-930. doi: 10.1016/j.cortex.2011.02.019

Moyer, A. (2013). *Foreign Accent: The Phenomenon of Non-Native Speech*. Cambridge, UK: Cambridge University Press.

Munson, B. (2011). The influence of actual and imputed talker gender on fricative perception, revisited (L). *The Journal of the Acoustical Society of America, 130*(5), 2631-2634. doi: 10.1121/1.3641410

Niedzielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology, 18*(1), 62-85.

Núñez-Peña, M. I., & Honrubia-Serrano, M. L. (2004). P600 related to rule violation in an arithmetic task. *Cognitive Brain Research, 18*(2), 130-141. doi: 10.1016/j.cogbrainres.2003.09.010

Nygaard, L. C. (2008). Perceptual integration of linguistic and nonlinguistic properties of speech. In D. B. Pisoni & R. E. Remez (Eds.), *The Handbook of Speech Perception* (pp. 390-413). Malden, MA: Blackwell Publishing Ltd.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97-113.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience, 2011*.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31*(6), 785-806. doi: 10.1016/0749-596X(92)90039-Z

Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language, 34*(6), 739-773. doi: 10.1006/jmla.1995.1033

Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience, 10*(6), 717-733. doi: 10.1162/089892998563121

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *The Journal of the Acoustical Society of America, 24*(2), 175-184. doi: 10.1121/1.1906875

Piai, V., Meyer, L., Schreuder, R., & Bastiaansen, M. C. M. (2013). Sit down and read on: Working memory and long-term memory in particle-verb processing. *Brain and Language, 127*(2), 296-306.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences, 36*(04), 329-347.

Qpzm Localstats Australia (2016). LocalStats and Demographics of Australia. Retrieved from http://localstats.com.au/

R Core Team (2014). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Rispens, J., & Been, P. (2007). Subject-verb agreement and phonological processing in developmental dyslexia and specific language impairment (SLI): A closer look. *International Journal of Language & Communication Disorders, 42*(3), 293-305.

Royle, P., Drury, J. E., & Steinhauer, K. (2013). ERPs and task effects in the auditory processing of gender agreement and semantics in French. *The Mental Lexicon, 8*(2), 216-244.

Severens, E., Jansma, B. M., & Hartsuiker, R. J. (2008). Morphophonological influences on the comprehension of subject–verb agreement: An ERP study. *Brain Research, 1228*, 135-144.

Swaab, T. Y., Ledoux, K., Camblin, C. C., & Boudewyn, M. A. (2012). Language-related ERP components. In S. J. Luck & E. S. Kappenman (Eds.), *The Oxford Handbook of Event-Related Potential Components* (pp. 397-439). New York: Oxford University Press.

Szewczyk, J. M., & Schriefers, H. (2011). Is animacy special?: ERP correlates of semantic violations and animacy violations in sentence processing. *Brain Research, 1368*, 208-221. doi: 10.1016/j.brainres.2010.10.070

Tanner, D., & Bulkes, N. Z. (2015). Cues, quantification, and agreement in language comprehension. *Psychonomic Bulletin & Review, 22*(6), 1753-1763. doi: 10.3758/s13423-015-0850-3

Tanner, D., & Van Hell, J. G. (2014). ERPs reveal individual differences in morphosyntactic processing. *Neuropsychologia, 56*, 289-301. doi: 10.1016/j.neuropsychologia.2014.02.002

Tesink, C. M. J. Y., Petersson, K. M., van Berkum, J. J. A., van den Brink, D., Buitelaar, J. K., & Hagoort, P. (2009). Unification of speaker and meaning in language comprehension: An fMRI study. *Journal of Cognitive Neuroscience, 21*(11), 2085-2099. doi: 10.1162/jocn.2008.21161

Ting, I., & Walters, C. (2016). Sydney's melting pot of language. *The Sydney Morning Herald*. Retrieved from http://www.smh.com.au/data-point/sydney-languages

Van Berkum, J. J. A., van den Brink, D., Tesink, C. M. J. Y., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience, 20*(4), 580-591. doi: 10.1162/jocn.2008.20054

van den Brink, D., Van Berkum, J. J. A., Bastiaansen, M. C. M., Tesink, C. M. J. Y., Kos, M., Buitelaar, J. K., & Hagoort, P. (2012). Empathy matters: ERP evidence for inter-individual differences in social language processing. *Social Cognitive and Affective Neuroscience, 7*(2), 173-183. doi: 10.1093/scan/nsq094

Van Engen, K. J., & Peelle, J. E. (2014). Listening effort and accented speech. *Frontiers in Human Neuroscience, 8*. doi: 10.3389/fnhum.2014.00577

van Herten, M., Chwilla, D. J., & Kolk, H. H. J. (2006). When heuristics clash with parsing routines: ERP evidence for conflict monitoring in sentence perception. *Journal of Cognitive Neuroscience, 18*(7), 1181-1197. doi: 10.1162/jocn.2006.18.7.1181

Van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology, 67*(6), 1176-1190.

van Petten, C., & Rheinfelder, H. (1995). Conceptual relationships between spoken words and environmental sounds: Event-related brain potential measures. *Neuropsychologia, 33*(4), 485-508. doi: 10.1016/0028-3932(94)00133-A

Vieru, B., Boula de Mareüil, P., & Adda-Decker, M. (2011). Characterisation and identification of non-native French accents. *Speech Communication, 53*(3), 292-310. doi: 10.1016/j.specom.2010.10.002

Wakabayashi, A., Baron-Cohen, S., Wheelwright, S., Goldenfeld, N., Delaney, J., Fine, D., . . . Weil, L. (2006). Development of short forms of the Empathy Quotient (EQ-Short) and the Systemizing Quotient (SQ-Short). *Personality and Individual Differences, 41*(5), 929-940. doi: 10.1016/j.paid.2006.03.017

Weber, A., Di Betta, A. M., & McQueen, J. M. (2014). Treack or trit: Adaptation to genuine and arbitrary foreign accents by monolingual and bilingual listeners. *Journal of Phonetics, 46*, 34-51. doi: 10.1016/j.wocn.2014.05.002

Wechsler, D., Coalson, D. L., & Raiford, S. E. (2008). *WAIS-IV: Wechsler adult intelligence scale*: Pearson San Antonio, TX.

Xu, N., & Demuth, K. (2012). Acoustic analysis of English codas by Mandarin learners of English. In F. Cox, K. Demuth, S. Lin, K. Miles, S. Palethorpe, J. Shaw & I. Yuen (Eds.), *Proceedings of the 14th Australasian International Conference on Speech Science and Technology* (pp. 41-44). Sydney, Australia: The Australasian Speech Science and Technology Association Incorporated.

Appendix A – Complete list of stimuli

Correct sentences are shown in the left column, with their corresponding incorrect versions on the right.

Morphologically correct	*Morphologically incorrect*
The mother carefully bakes.	The mothers carefully bakes.
The mothers carefully bake.	The mother carefully bake.
The puppy happily barks.	The puppies happily barks.
The puppies happily bark.	The puppy happily bark.
The captain typically bats.	The captains typically bats.
The captains typically bat.	The captain typically bat.
The tiger typically bites.	The tigers typically bites.
The tigers typically bite.	The tiger typically bite.
The baby typically burps.	The babies typically burps.
The babies typically burp.	The baby typically burp.
The painter happily chats.	The painters happily chats.
The painters happily chat.	The painter happily chat.
The teacher carefully checks.	The teachers carefully checks.
The teachers carefully check.	The teacher carefully check.
The father carefully cooks.	The fathers carefully cooks.
The fathers carefully cook.	The father carefully cook.
The builder carefully cuts.	The builders carefully cuts.
The builders carefully cut.	The builder carefully cut.
The neighbour typically fights.	The neighbours typically fights.
The neighbours typically fight.	The neighbour typically fight.
The neighbour happily hikes.	The neighbours happily hikes.

The neighbours happily hike.	The neighbour happily hike.
The lady carefully knits.	The ladies carefully knits.
The ladies carefully knit.	The lady carefully knit.
The caller carefully knocks.	The callers carefully knocks.
The callers carefully knock.	The caller carefully knock.
The athlete carefully leaps.	The athletes carefully leaps.
The athletes carefully leap.	The athlete carefully leap.
The kitten happily licks.	The kittens happily licks.
The kittens happily lick.	The kitten happily lick.
The uncle typically naps.	The uncles typically naps.
The uncles typically nap.	The uncle typically nap.
The dancer happily shakes.	The dancers happily shakes.
The dancers happily shake.	The dancer happily shake.
The hunter carefully shoots.	The hunters carefully shoots.
The hunters carefully shoot.	The hunter carefully shoot.
The sister happily shops.	The sisters happily shops.
The sisters happily shop.	The sister happily shop.
The leader typically shouts.	The leaders typically shouts.
The leaders typically shout.	The leader typically shout.
The pilot typically sits.	The pilots typically sits.
The pilots typically sit.	The pilot typically sit.
The daughter happily talks.	The daughters happily talks.
The daughters happily talk.	The daughter happily talk.
The author happily types.	The authors happily types.

The authors happily type.	The author happily type.
The baker happily waits.	The bakers happily waits.
The bakers happily wait.	The baker happily wait.
The learner carefully writes.	The learners carefully writes.
The learners carefully write.	The learner carefully write.

Semantically correct	*Semantically incorrect*
The lady bites the cupcake.	The lady bites the suitcase.
The barman cools the cocktail.	The barman cools the carpet.
The helper cuts the carrot.	The helper cuts the buses.
The mother feeds the baby.	The mother feeds the doorbell.
The butler fills the teapot.	The butler fills the carrot.
The builders hang the picture.	The builders hang the salad.
The doctor heals the soldier.	The doctor heals the cottage.
The fighters hit the target.	The fighters hit the daylight.
The captain leads the soldiers.	The captain leads the teacup.
The actor leaves the party.	The actor leaves the ceiling.
The farmers light the torches.	The farmers light the cabbage.
The owners lock the cottage.	The owners lock the dinner.
The teachers mark the pages.	The teachers mark the summer.
The sailor meets the captain.	The sailor meets the painting.
The father moves the table.	The father moves the sunshine.
The nanny packs the suitcase.	The nanny packs the colour.
The teachers park the buses.	The teachers park the pages.

The uncle rides the camel.	The uncle rides the season.
The owner rolls the carpet.	The owner rolls the birthday.
The sisters save the kitten.	The sisters save the puddle.
The doctor signs the papers.	The doctor signs the curry.
The nurses soak the bandage.	The nurses soak the cocktail.
The sisters toss the salad.	The sisters toss the siren.
The daughters wash the dishes.	The daughters wash the danger.
The women wipe the table.	The women wipe the concert.

Native speaker's fillers

She wore a woolly jumper.

I'd like to meet a penguin.

His friend was from Japan.

My mum went away last week.

My cousin broke his arm.

The girl likes to eat cake.

Make sure you water the plants.

I'd like to visit China.

She is scared of crocodiles.

Her house is across the road.

His brother is an actor.

He supports the football team.

She wouldn't lend me her watch.

My favourite one is red.

Foreign-accented speaker's fillers

I went shopping yesterday.

Teachers often eat donuts.

The cat sat on the chair.

The mug was full of tea.

Coffee is her favourite.

He always wears a hat.

Empty space is useful.

Birds fly over the mountain.

The house is falling down.

My aunt has bought a new house.

The library is closed.

The child learns how to read.

She heard some loud noises.

Rabbits are hard to find.

The vet has a pet hamster.

I heard he fought a bear.

The rollercoaster is great.

The test gave me a headache.

The cleaner liked to dance.

Daisies are often yellow.

She thought the paint had dried.

I met a weightlifter.

There were too many biscuits.

Her hobby is rock-climbing.

The parrot learned to talk.

I kept a souvenir.

I love to watch the clouds.

The bus has broken down.

My grandmother has gone deaf.

Her phone rang during class.

I kept it in the cabinet.

The lecturer lost her voice.

We went on a picnic.

My dad built a treehouse.

The test is tomorrow.

Luckily I packed a snack.

We all decided to leave.

She tripped over a step.

Travel can be confronting.

Africa is far away.

The boat sank under the waves.

Always use a sharp pencil.

The sky was clear last night.

Peppermint tea is nice.

Her birthday is in July.

I will be on holiday.

The shopping bags were heavy.

I asked but he didn't know.

No one was going to sleep.

She sometimes loses hair ties.

Everyone went crazy.

The judge filled in the details.

Ferrets are always grumpy.

She's moving to Canada.

Kites fly higher than swans.

The semester seemed long.

There were many useful books.

I made a chocolate cake.

The apple fell from the tree.

She had a hula hoop.

The accountant was wrong.

The headlights were very bright.

The singer was nervous.

They dressed as superheroes.

They thought she had forgotten.

The spelling was correct.

He left the book behind.

The stapler had no staples.

This is the best café.

She thought she failed the subject.

Scones are best with jam and cream.

My train is running late.

I think it will snow today.

Monkeys like to live in trees.

He was sick yesterday.

I'll be back tomorrow.

The kangaroo was hungry.

She decided on France.

They ran out of water.

The bathtub was empty.

They had pink and red flowers.

The professor spilled the milk.

Your scarf is brightly coloured.

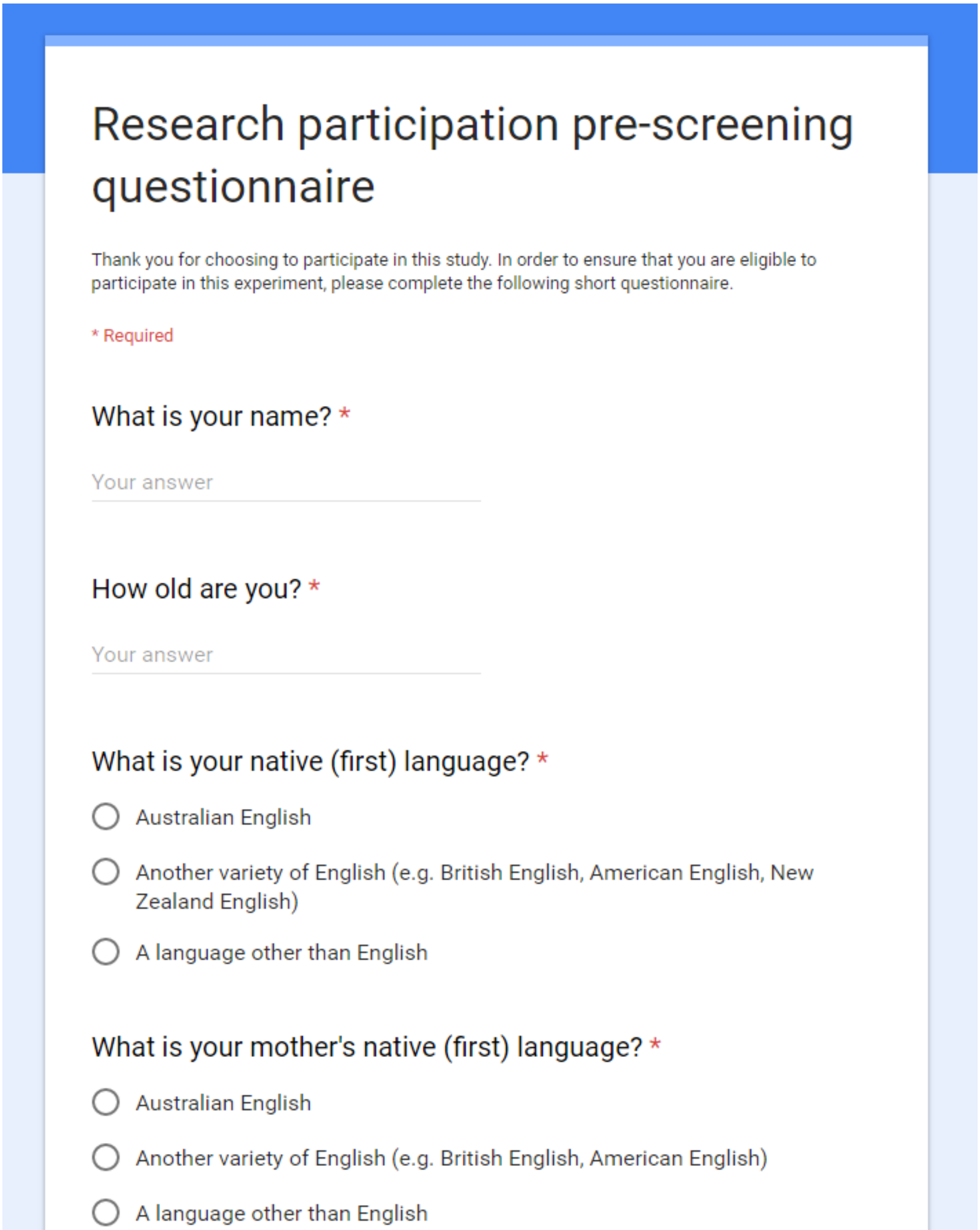
The porridge had gone cold.

The orange one was the best.

His punctuation was poor.

Appendix B – Screening questionnaire

All participants completed this online questionnaire before attending the test session.



Research participation pre-screening questionnaire

Thank you for choosing to participate in this study. In order to ensure that you are eligible to participate in this experiment, please complete the following short questionnaire.

* Required

What is your name? *

Your answer

How old are you? *

Your answer

What is your native (first) language? *

- Australian English
- Another variety of English (e.g. British English, American English, New Zealand English)
- A language other than English

What is your mother's native (first) language? *

- Australian English
- Another variety of English (e.g. British English, American English)
- A language other than English

What is your father's native (first) language? *

- ◯ Australian English
- ◯ Another variety of English (e.g. British English, American English)
- ◯ A language other than English

Do you speak any language other than English? *

- ◯ Yes, I speak a language other than English fluently
- ◯ I have studied another language (e.g. at school or university), but I am not a fluent speaker
- ◯ No, I speak English only

Are you right handed? *

- ◯ Yes
- ◯ No

Have you ever been diagnosed with any hearing, language, neurological or cognitive impairment? *

- ◯ Yes
- ◯ No

If you answered 'yes' to the above question, please provide details.

Your answer

SUBMIT

Appendix C – Pre-task questionnaire

During the test session, all participants completed the questionnaire reproduced on the following pages before completing the EEG task.

Department of Linguistics and Centre for Language Sciences
Faculty of Human Sciences
MACQUARIE UNIVERSITY NSW 2109

Phone: +61 2 9850 8783
Fax: +61 2 9850 9199
Email: ling.cll@mq.edu.au

Participant number (experimenter to complete): ____________________

Project: Processing of Spoken Language

Participant Questionnaire

Name: ____ Age: __________ Gender: ____________

Thank you for participating in this research study. To assist in our data analysis, please answer the following questions as accurately as possible. If you have any questions please ask the experimenter.

1. Is English your first language? Yes No_

   If you answered 'no', what is your first language? ____________________________

2. Do you speak any languages other than English? Yes No_

   If you answered 'yes', please list below the language(s) you speak, the approximate frequency that you use the language (please choose from the following: daily, weekly, monthly, rarely (i.e. less than once per month), never) and an indication of your proficiency (please choose from the following: beginner, intermediate learner, advanced learner, proficient, native speaker). For example: *Spanish, twice per week, intermediate learner; French, rarely, beginner.*

   ____

   ____

   ____

   ____

3. Were your parents born outside Australia? Yes No

If you answered 'yes', please list below their country of birth. For example: *Mother: New Zealand, Father: England*

____

4. Have you been diagnosed with any hearing, language, neurological or cognitive impairments? Yes No

If you answered 'yes', please provide a brief description below.

____

____

____

____

5. Please indicate in the table below which hand you would normally use to complete each activity. Place a cross ('x') in the relevant column for each activity. Where your preference is so strong that you would never try to use the other hand unless absolutely forced to, place two crosses in the relevant box. If for any activity you are really indifferent about which hand you use, please place one cross in each box (left and right). Please only leave a row blank if you have no experience of completing that particular task.

Task	**Left**	**Right**
Writing		
Drawing		
Throwing		
Using scissors		
Using a toothbrush		
Using a knife (without a fork)		
Using a spoon		
Using a sports racquet (e.g. tennis, badminton)		
Using a computer mouse		
Removing the lid of a box		
Brushing/combing hair		
Using a hammer		
Using a key to unlock a door		
Which foot would you normally kick with?		
Which eye do you use when you only use one?		

6. What is the highest level of education that you have completed? For example: *HSC* or *Bachelor's degree*

____

7. Have you lived in any country other than Australia for a period of more than six months?

Yes No

If you answered 'yes', please list below all countries you have lived in, the duration of your stay and your age at that time. For example: *New Zealand, 3 years, age 12-15*

____

____

____

8. What suburb do you live in? ____

9. Please list any other suburbs you have lived in in the past (within Australia). If the suburb is outside of Sydney, please also indicate the state.

____

____

____

10. Do you regularly interact with people who speak English with a foreign accent (i.e. not an Australian accent)?

Yes No

If you answered 'yes', please list below the accents to which you are regularly exposed, the approximate frequency of exposure (please choose from the following: daily, weekly, monthly, rarely (i.e. less than once per month), never) and the context in which this exposure occurs. For example: *Indian, daily, work colleagues; Italian, monthly, friend; Chinese, weekly, social group.*

____

____

____

____

____

____

11. Please complete the following table by circling the number that corresponds to the response best representing you for each statement.

	Strongly agree	Slightly agree	Slightly disagree	Strongly disagree
I can easily tell if someone else wants to enter a conversation.	1	2	3	4
I really enjoy caring for other people.	1	2	3	4
I find it hard to know what to do in a social situation.	1	2	3	4
I often find it difficult to judge if something is rude or polite.	1	2	3	4
In a conversation, I tend to focus on my own thoughts rather than on what my listener might be thinking.	1	2	3	4
I can pick up quickly if someone says one thing but means another.	1	2	3	4
It is hard for me to see why some things upset people so much.	1	2	3	4
I find it easy to put myself in somebody else's shoes.	1	2	3	4
I am good at predicting how someone will feel.	1	2	3	4
I am quick to spot when someone in a group is feeling awkward or uncomfortable.	1	2	3	4
I can't always see why someone should have felt offended by a remark.	1	2	3	4
I don't tend to find social situations confusing.	1	2	3	4
Other people tell me I am good at understanding how they are feeling and what they are thinking.	1	2	3	4
I can easily tell if someone else is interested or bored with what I am saying.	1	2	3	4
Friends usually talk to me about their problems as they say that I am very understanding.	1	2	3	4
I can sense if I am intruding, even if the other person doesn't tell me.	1	2	3	4
Other people often say that I am insensitive, though I don't always see why.	1	2	3	4
I can tune in to how someone feels rapidly and intuitively.	1	2	3	4
I can easily work out what another person might want to talk about.	1	2	3	4
I can tell if someone is masking their true emotion.	1	2	3	4
I am good at predicting what someone will do.	1	2	3	4
I tend to get emotionally involved with a friend's problems.	1	2	3	4

This is the end of the questionnaire. Thank you for your participation.

Appendix D – Post-task questionnaire

During the test session, all participants completed the questionnaire reproduced on the following page after completing the EEG task.

MACQUARIE University

Department of Linguistics and Centre for Language Sciences
Faculty of Human Sciences
MACQUARIE UNIVERSITY NSW 2109

Phone: +61 2 9850 8783
Fax: +61 2 9850 9199
Email: ling.cll@mq.edu.au

Participant number (experimenter to complete): ___________________

Project: Processing of Spoken Language

Post-task Questionnaire

Thank you for participating in this research study. Having now completed the EEG task, please answer the following questions. If you have any questions please ask the experimenter.

1. During the EEG task, you heard sentences spoken by two speakers. One was a native speaker of Australian English, the other had a foreign accent. What nationality do you think the foreign-accented speaker was?

   ____

2. Please rate how strong you thought the foreign-accented speaker's accent was by circling the corresponding number on the following scale, where one indicates no foreign accent and seven indicates a very strong foreign accent.

1	2	3	4	5	6	7
No foreign accent						Very strong foreign accent

This is the end of the questionnaire. Thank you for your participation.

Appendix E – Details of additional analyses

Analysis without baseline correction

A cluster-based permutation test was conducted on data which had been pre-processed following all specifications set out in the methodology (section 3.4) but without use of a baseline correction. The purpose of this was to determine that no significant differences between conditions occurred during the proposed baseline period. This analysis found that, while no significant clusters occurred during the proposed baseline period for the semantic and omission error conditions, two significant negative clusters (0-79 ms and 150-312 ms) were observed during the proposed baseline period for the commission error condition (the proposed baseline period was 0-200 ms).

Analysis for first and second halves of the data

After conducting the cluster-based permutation tests on the morphosyntactic-condition data reported in section 4.1, further cluster-based permutation tests were conducted as a follow-up. The dataset was divided in half (i.e. responses to the first half of the trials in each condition formed one dataset and responses to the second half of the trials in each condition formed the other) and cluster-based permutation tests were performed on each half separately. The purpose of this was to look for any effects which may be significant in the first half of the data, but not appear in the second half due to participants' acclimatisation to the violations in the stimuli. This followed the procedure employed in Hanulikovà and colleagues' study (2012), in which it was found that the significant P600 effects observed were entirely driven by responses in the first half of the experiment, and no significant grammaticality effects were observed in the second half of the data. This proved not to be the case in the current study, as the clusters identified in the separate cluster-based permutation tests on the two halves of the data were almost identical both to each other and to the cluster observed in the analysis of the complete dataset. Therefore, data from only the time window identified in the initial cluster-based permutation test, 210 to 330ms, was entered into the ANOVA.

Office of the Deputy Vice-Chancellor
(Research)

Research Office
Research Hub, Building C5C East
Macquarie University
NSW 2109 Australia
T: +61 (2) 9850 4459
http://www.research.mq.edu.au/
ABN 90 952 801 237

MACQUARIE University
SYDNEY·AUSTRALIA

2 March 2016

Professor Katherine Demuth
Department of Linguistics
Faculty of Human Sciences
Macquarie University
NSW 2109

Dear Professor Demuth

Reference No: 5201600062

Title: *Neural Responses to Morphosyntactic Violations in Foreign-Accented Speech*

Thank you for submitting the above application for ethical and scientific review. Your application was considered by the Macquarie University Human Research Ethics Committee (HREC (Human Sciences & Humanities)) at its meeting on 26 February 2016.

I am pleased to advise that ethical and scientific approval has been granted for this project to be conducted at:

- Macquarie University

This research meets the requirements set out in the *National Statement on Ethical Conduct in Human Research* (2007 – Updated May 2015) (the *National Statement*).

This letter constitutes ethical and scientific approval only.

Standard Conditions of Approval:

1. Continuing compliance with the requirements of the *National Statement,* which is available at the following website:

http://www.nhmrc.gov.au/book/national-statement-ethical-conduct-human-research

2. This approval is valid for five (5) years, subject to the submission of annual reports. Please submit your reports on the anniversary of the approval for this protocol.

3. All adverse events, including events which might affect the continued ethical and scientific acceptability of the project, must be reported to the HREC within 72 hours.

4. Proposed changes to the protocol must be submitted to the Committee for approval before implementation.

It is the responsibility of the Chief investigator to retain a copy of all documentation related to this project and to forward a copy of this approval letter to all personnel listed on the project.

Should you have any queries regarding your project, please contact the Ethics Secretariat on 9850 4194 or by email ethics.secretariat@mq.edu.au

The HREC (Human Sciences and Humanities) Terms of Reference and Standard Operating Procedures are available from the Research Office website at:

http://www.research.mq.edu.au/for/researchers/how_to_obtain_ethics_approval/human_research_ethics

The HREC (Human Sciences and Humanities) wishes you every success in your research.

Yours sincerely

Dr Karolyn White
Director, Research Ethics & Integrity,
Chair, Human Research Ethics Committee (Human Sciences and Humanities)

This HREC is constituted and operates in accordance with the National Health and Medical Research Council's (NHMRC) *National Statement on Ethical Conduct in Human Research* (2007) and the *CPMP/ICH Note for Guidance on Good Clinical Practice*.

Details of this approval are as follows:

Approval Date: 26 February 2016

The following documentation has been reviewed and approved by the HREC (Human Sciences & Humanities):

Documents reviewed	Version no.	Date
Macquarie University Ethics Application Form		Received 9/2/2016
Recruitment Advertisement	1	9/2/2016
Macquarie University Participant Information and Consent Form (PICF) entitled *'Processing of Foreign-Accented Speech'*	1	6/2/2016
Information Sheet: Electroencephalography (EEG) Study	1	9/2/2016
Participant Questionnaire	1	5/2/2016

Ethics amendment Ref. 5201600062

APPROVED
By Fran Thorp at 2:00 pm, Jul 26, 2016

Fran Thorp

Rebecca Holt

Tue 26/07/2016 11:05 AM

To:Ethics Secretariat <ethics.secretariat@mq.edu.au>;

Cc:Katherine Demuth <katherine.demuth@mq.edu.au>;

2 attachments (718 KB)

ethics amendment R Holt 26-07-16.pdf; Questionnaire 16_07_26.docx;

Hello,
Please find attached another (small) ethics amendment.
Thank you,
Rebecca

Human Research Ethics Committee

REQUEST FOR AMENDMENT FORM

Please complete this form for all amendments/modifications including extensions to approved ethics projects.

For quick and efficient review of your amendment, please provide sufficient information in this document to allow the amendment to be reviewed as a standalone document (i.e. it does not require the Ethics Secretariat or HREC reviewing the original application).

Please attach tracked and clean copies of all amended documents to the amendment request. Documents could include participant information and consent forms (PICF), advertising material, surveys, interview questions, verbal scripts, support letters from external organizations.

Submitting this form:

HREC approved applications: Please send this form to ethics.secretariat@mq.edu.au.

Faculty/School-approved applications:
Please send this form to the ethics subcommittee administrator of the relevant Faculty/School
Faculty of Human Sciences: fhs.ethics@mq.edu.au
Faculty of Science and Engineering: sci.ethics@mq.edu.au
Faculty of Arts: artsro@mq.edu.au
Faculty of Business and Economics: fbe-ethics@mq.edu.au
MGSM: ethics@mgsm.edu.au
PACE: pace.ethics@mq.edu.au
Faculty of Medicine and Health Sciences: ethics.secretariat@mq.edu.au.

Handwritten forms will not be accepted.

---

1. **Human Research Ethics Committee Reference No:** 5201600062

2. **Chief Investigator/Supervisor:** Professor Katherine Demuth

   Faculty: Human Sciences

   Department: Linguistics

   Email: katherine.demuth@mq.edu.au

   Date of amendment: 26/07/16