Highlights

- Both nonnative accent and multi-talker background noise impact online lexical-semantic access, indexed by both the N400 magnitude and onset latency.
- Listening to sentences presented in noise is associated with higher listening effort as indexed by an increase in alpha power activity.
- Semantic anomalies presented in multi-talker background noise elicit higher theta activity, whereas processing nonnative-accented anomalies is associated with decreased theta activity.
- The undesirable effect of multi-talker background noise is augmented when listening to nonnative-accented speech compared to native accented speech, in behavioral comprehension accuracy measures.

Abstract

We examined the neural correlates underlying the semantic processing of native- and nonnative-accented sentences, presented in quiet or embedded in multi-talker noise. Implementing a semantic violation paradigm, 36 English monolingual young adults listened to American-accented (native) and Chinese-accented (nonnative) English sentences with or without semantic anomalies, presented in quiet or embedded in multi-talker noise, while EEG was recorded. After hearing each sentence, participants verbally repeated the sentence, which was coded and scored as an offline comprehension accuracy measure. In line with earlier behavioral studies, the negative impact of background noise on sentence repetition accuracy was higher for nonnative-accented than for native-accented sentences. At the neural level, the N400 effect for semantic anomaly was larger for native-accented than for nonnative-accented sentences, and was also larger for sentences presented in quiet than in noise, indicating impaired lexical-semantic access when listening to nonnative-accented speech or sentences embedded in noise. No semantic N400 effect was observed for nonnative-accented sentences presented in noise. Furthermore, the frequency of neural oscillations in the alpha frequency band (an index of online cognitive listening effort) was higher when listening to sentences in noise versus in quiet, but no difference was observed across the accent conditions. Semantic anomalies presented in background noise also elicited higher theta activity, whereas processing nonnative-accented anomalies was associated with decreased theta activity. Taken together, we found that listening to nonnative accents or background noise is associated with processing challenges during online semantic access, leading to decreased comprehension accuracy. However, the underlying cognitive mechanism (e.g., associated listening efforts) might manifest differently across accented speech processing and speech in noise processing.

Keywords: nonnative accented speech, speech in noise, EEG, semantic access

Introduction

In today's globalized world, people increasingly encounter speakers with a nonnative accent: the majority of English speakers, for example, are nonnative speakers (Eberhard et al., 2022). Moreover, natural speech communication rarely takes place in ideal listening conditions, as we often listen to others when surrounded by background noise. Processing and comprehending nonnative-accented speech in noisy backgrounds has thus become a common characteristic of everyday communication in our multicultural and multilingual society. However, few if any neurolinguistic studies have examined how listeners process nonnative-accented sentences in background noise. Therefore, in the current study, we examined the neural correlates underlying the processing of nonnative-accented and native-accented sentences embedded in background speech noise and in quiet conditions, employing electrophysiological techniques (Event Related Potentials and time-frequency analyses).

Nonnative-accented speech processing

Nonnative-accented speech can be defined as "non-pathological speech that differs in some noticeable respects from native speaker pronunciation norms" (Munro & Derwing, 1995, page 298). Most researchers have rejected the notion that the L1 and L2 phonetic subsystems of bilinguals are completely separable from each other (Grosjean, 1998; Lambert & Rawlings, 1969). For example, the Speech Learning Model (Flege et al., 1999, 2003; Flege & Bohn, 2021) proposes that the L1 and L2 phonetic subsystems in bilinguals exist in a "common phonological space", and interact with each other, resulting in perceived nonnative accents. More specifically, because of L1 phonetic interference, nonnative-accented L2 speech production often contains phonological approximations (for a review, see Nagle & Baese-Berk, 2022), such as substituting phonemes in L2 by similar native L1 sounds (Huang & Evanini, 2016; Wester et al., 2007).

Nonnative-accented speech can thus pose comprehension difficulty to native ears, as listeners must reconcile incoming deviating acoustic signals with their existing native phonetic categories. According to the classic abstract-entry model (McClelland & Elman, 1986; Norris, 1999; Pallier et al., 2001) this mapping between the acoustic signal and abstract phonological representations occurs at a pre-lexical processing stage (Samuel & Kraljic, 2009; Sumner, 2011). Specifically, with greater perceptual distance between the incoming acoustic signal and listeners' L1-based phonological representations, which typifies nonnative-accented speech, it is proposed that the activation of the target lexical items and higher-order lexical-semantic access might require more top-down lexical involvement (Goslin et al., 2012) and greater cognitive resource recruitment (Van Engen & Peelle, 2014), resulting in impaired comprehension.

Rather than examining smaller linguistic unit processing (e.g., phoneme perception, see Binder, 2016, for a review), the focus of this study is on sentence-level processing. Behavioral research on nonnative-accented sentence processing has generally found that, relative to native-accented sentences, nonnative-accented sentence processing is associated with certain processing costs, as indexed by a slower processing rate and lower comprehension accuracy. Such costs were found across a wide range of tasks, including sentence transcription (Gordon-Salant et al., 2010), true-or-false sentence judgment (Munro & Derwing, 1995), and visual probe-word judgment (Clarke & Garrett, 2004).

In addition to studies using behavioral measurements, recent studies have used electroencephalogram (EEG) techniques to examine online neural processing of nonnative-accented speech (e.g., Caffarra & Martin, 2019; Goslin et al., 2012; Grey & Van Hell, 2017;

Hanulíková et al., 2012; Holt et al., 2018; Romero-Rivas et al., 2015, 2016; for a review, see Caffarra et al., 2023). The EEG technique provides a noninvasive measurement of scalp-recorded electrical activity in the order of milliseconds. Event Related Potentials (ERPs) reflect EEG activity time-locked to certain events (e.g., sensory, motor or cognitive) in order to capture neural changes during neurocognitive processing as it unfolds over time. Furthermore, time-frequency analysis extracts time frequency representations (TFRs) from EEG signals, such as alpha, beta, and gamma waves. ERPs/TFRs enable the investigation of real-time temporal dynamics as listeners process accented speech.

A widely employed paradigm in EEG sentence processing experiments is the semantic-violation paradigm in which participants read or listen to sentences containing target words that are either congruent (i.e., semantically well-formed) or incongruent (i.e., semantically anomalous) with the preceding semantic context (Behrman & Akhund, 2013). Semantically anomalous words embedded in sentences typically elicit the N400 response, an electrophysiological marker of lexical-semantic access (Kutas & Federmeier, 2011; Federmeier, 2021). The N400 component is a negative-going ERP component that peaks around 400 ms post-stimulus onset. When the target word meets the expectation generated by a preceding context, such as in semantically well-formed sentences (e.g., the target word "butter" in "He spread the warm bread with butter"; Kutas & Hillyard, 1980), the N400 amplitude is less negative in comparison to when the target word is anomalous in the context (e.g., the target word "socks" in "He spread the warm bread with socks"; Kutas & Hillyard, 1980). Because of high sensitivity to preceding semantic context, the N400 effect is regarded as a reliable index for semantic access.

Hanulíková et al. (2012) is among the first ERP studies that examine nonnative-accented sentence processing in native listeners. Even though the primary focus of the study was on syntactic processing (not further discussed here), sentences with or without semantic violations were also included (e.g., "It was very cold last night, so I put a thick blanket/*evening on my bed"). Testing native Dutch listeners, Hanulíková et al. (2012) found that the N400 effect was more broadly distributed for Turkish-accented Dutch sentences than for native Dutch-accented sentences. However, the magnitude of the N400 effect did not differ across the accent conditions and the authors concluded that the N400 effect was not modulated by the speakers' accents.

Adopting a similar design but including four nonnative accents (French-, Greek-, Italian-, and Japanese-accented Spanish), Romero-Rivas et al. (2015) tested native Spanish listeners and replicated the finding that the N400 effect was more widely distributed for the nonnative-accented than for the native-accented Spanish sentences. The N400 response associated with nonnative-accented sentences was also more negative-going than the N400 associated with native-accented sentences. According to the authors, this enhanced N400 emerged because the semantic violation was harder to process in nonnative-accented sentences than in native-accented sentences. Additionally, following the N400 effect, semantic anomalous sentences produced with a native accent also elicited a P600 re-analysis effect, while this effect was absent during nonnative-accented sentence processing. The authors took these findings to imply that listeners did not seek an alternative resolution to re-analyze the semantic violation when the sentence was produced by nonnative-accented speakers.

Grey and Van Hell (2017) and Holt et al. (2018) also implemented the semantic violation paradigm in their ERP experiments and tested native English listeners' processing of Mandarin Chinese-accented English sentences. In Grey and Van Hell (2017), all participants were monolinguals with little experience listening to nonnative-accented speech, and in the post-experiment survey only 36% correctly identified the nonnative accent as 'Asian'. In comparison, the majority of participants in Holt et al. (2018) had greater experience with nonnative-accented speech, and 63% correctly identified the nonnative accent. In both studies, semantic anomalies in native-accented sentences elicited the N400 effect. In contrast, semantic anomalies in Chinese-accented sentences elicited no N400 effect in the typical time-window, but a late negativity effect in the later time-window in Grey and Van Hell (2017). However, Holt et al. (2018) observed no difference in N400 effects (in terms of topographical distribution, onset latency, magnitude and latency) for semantic anomalies in native- versus nonnative-accented sentences.

More recently, testing how Dutch native listeners processed nonnative-accented Dutch sentences (four speakers: one German, one Southern-Irish English, one Macedonian, and one Polish), Schiller et al. (2020) observed a delayed negativity (400 - 600 ms) with a narrower time-window for unexpected words produced by nonnative-accented speakers compared to native speakers. This delayed or extended N400 effect was also reported in two other studies: when American-English listeners heard semantically incongruent words in Spanish-accented English sentences (Wambacq et al., 2023), and when native Spanish listeners heard semantic anomalies produced by Chinese-accented speakers (Gosselin et al., 2021).

Taken together, prior ERP investigations found that the processing of semantic anomalies differs for native- and nonnative-accented sentences, demonstrated by differences in the N400 topographical distribution (Hanulíková et al., 2012), magnitude (Romero-Rivas et al., 2015), and onset latency (Gosselin et al., 2021; Grey & Van Hell, 2017; Schiller et al., 2020; Wambacq et al., 2023; for a review, see Caffarra et al., 2023). Of note is that all previous ERP studies presented the nonnative- and native-accented sentences in an ideal listening environment without any background noise, even though most real-life listening conditions are noisy. We currently do not know to what extent neurophysiological processing of nonnative- and native-accented speech is differentially affected by background noise, the question that the present study addressed.

Above, we have discussed the N400 as the main index of semantic processing in research on accented speech processing. EEG research using time frequency analysis has also reported that semantic processing is associated with changes in theta frequency (4 – 8 Hz), albeit it is not as extensively studied as the N400 effect (and, to the best of our knowledge, has not been studied in the context of accented speech processing). In an early study, Hagoort et al. (2004) reported a frontal midline theta power increase when participants read semantically anomalous words in sentences, and the effect overlapped with the N400 time-window. Such theta power increase associated with semantic violation was also reported in later studies on both sentence reading and listening, albeit the topographical distribution of the theta power increase varied. Specifically, theta power increase associated with reading semantic anomalies was found in mid-frontal (Hald et al., 2006), in bilateral posterior and left anterior (Davidson & Indefrey, 2007), and right posterior (Wang, Zhu, et al., 2012) regions. Other studies reported no theta power changes following reading semantic violations (Penolazzi et al., 2009; Rommers et al., 2013). In listening comprehension, theta power changes associated with semantic violations remain relatively underreported: Schneider and Maguire (2019) are among the few who reported a right posterior distribution for both adults and eight- to nine-year-old children. In comparison, Wang, Jensen et al. (2012) only reported an increased activity in the beta band when listening to semantically incongruent words, but not in the theta band.

So far, earlier studies that reported increased theta power associated with the processing of semantic violations only focused on language processing during ideal reading or listening conditions, without taking different accents or background noise into account. The present study therefore explores the extent to which accent and/or background noise incur changes in theta power associated with semantic processing during sentence comprehension (see Data analytic plan in the Methods section for details).

In the next section, we review EEG studies on semantic processing in speech in noise (that exclusively presented speech produced by native speakers).

Speech in noise processing

In the speech in noise literature, noise level is usually represented by the difference in decibels (dB) between the intensity of the target speech signal and the intensity of the background noise, which is often referred to as the signal-to-noise ratio (SNR). For instance, if the target speech is presented at 70 dB and the background noise is presented at 65 dB, the SNR would be +5 dB.

Despite the long history of empirical work on speech in noise processing, most of these studies focused on the neural correlates of attention allocation and inhibition control in a noisy background, and did not take semantic processing into account. The closest EEG study that examined sentential semantic processing focused on noise vocoded speech (Obleser & Kotz, 2011). Noise vocoded speech was originally developed to simulate the speech signal transduced by a cochlear implant (Shannon et al., 1995). Obleser and Kotz (2011) had participants listen to German noise vocoded sentences at three different levels (1-band, 4-band, 16-band). Cloze probability of the sentences was either low (e.g., “Sie wiegt das Mehl” [she weighs the flour]) or high (e.g., “Sie siebt das Mehl” [she sifts the flour]). It was found that the N400 effect was more sensitive to sentential cloze-probability when speech quality increased. Specifically, the N400 effect was smaller in magnitude and delayed in peak latency in the 4-band condition, compared to the least degraded condition (i.e., 16-band). At the most degraded 1-band condition the N400 effect was completely absent. Similar results were reported in Strauß et al. (2013), using a similar design with three degradation levels: 4-band, 8-band, and clear speech.

Noise vocoded speech is one type of degraded speech, but it is different from everyday background noise. In the current study, we used multi-talker babble as a type of background noise that is more representative of noisy environments in real-life situations, such as school cafeterias, classrooms, and cocktail parties. Although several recent EEG studies have examined speech processing in a multi-talker babble background, they focused on the usage of hearing aids with various levels of SNRs (Fiedler et al., 2021; Seifi Ala et al., 2020), but not on semantic processing.

Nonnative-accented speech in noise processing

Processing sentences in a noisy environment challenges the auditory system, and potentially exacerbates listeners' difficulties with processing nonnative-accented sentences relative to native-accented sentences. However, to our knowledge there are no published neurolinguistic studies that have examined neural processes associated with listening to nonnative-accented sentences in background noise. We therefore provide a brief overview of findings from behavioral research.

Behavioral research on nonnative-accented speech in noise processing in young adults has investigated word-level (Bent, 2014; Bent & Holt, 2013), phrase-level (McLaughlin et al., 2018) and sentence-level (Bent & Atagi, 2017; Gordon-Salant et al., 2013; Rogers et al., 2004; Wilson & Spaulding, 2010) processing. It is generally found that the undesirable effect of background noise is augmented when listening to nonnative-accented speech compared to native accented speech. This was found across a wide variety of tasks, such as true-or-false judgment (Wilson & Spaulding, 2010), sentence transcription (Rogers et al., 2004), and verbal repetition (Bent, 2018; Bent & Atagi, 2015).

To explain this much stronger negative effect of noise on the processing of nonnative-accented speech relative to native-accented speech, it has been proposed that because processing nonnative-accented speech is already more effortful and requires more attentional resources, the addition of background noise might leave the listeners without sufficient cognitive resources (Bent & Atagi, 2015; Rogers et al., 2004), which results in a significant decline in behavioral performance.

Higher demands of cognitive resources during processing

This emphasis on cognitive resources is also captured in Van Engen and Peelle's (2014) model of speech comprehension, proposing that different degrees of acoustic mismatch will require varying levels of cognitive recruitment. Specifically, when speech signals match listeners' perceptual expectations, speech processing is relatively automatic and comprehension can proceed with minimal effort. However, when incoming speech mismatches the listeners' perceptual expectations, for example when encountering unfamiliar accents and/or background noise, listeners must recruit additional cognitive resources to compensate. In extreme conditions with high degree of acoustic mismatch, full compensation from recruiting additional cognitive resources is impossible, which renders comprehension to be impossible as well.

A more direct measurement of listening effort during online native- and nonnative-accented speech in noise processing using EEG potentially sheds more light on the roles of nonnative accent and of background noise. In prior EEG studies, enhanced power in the alpha frequency band (8-12 Hz) has generally been linked to increased working memory load, inhibitory control, and selective attention (Foxe & Snyder, 2011; Hunter, 2020; Tuladhar et al., 2007). For instance, in a modified Sternberg task, participants were presented with a memory set which consisted of two, four, or six letters (Jensen et al., 2002). After the screen went blank, participants mentally retained the memorization set for 2.8 seconds and when the probe appeared, they judged whether the probe letter belonged to the memorization set. Power spectral analysis revealed that activity in the alpha frequency band in the parietal region increased with more letters in the memory set, indexing higher working memory load.

Turning to sentence processing, parietal alpha power has been found to increase when participants mentally held an argument (i.e., subject or object) over longer distances (Meyer et al., 2013). An increase in oscillatory power in the alpha frequency band has also been associated with attentive listening to sentences as they unfold over time (Peña & Melloni, 2012) and with listening to sentences that switch from bilingual listeners' weaker language to the dominant language (Fernandez et al., 2019).

With respect to acoustic degradation, alpha power has been found to be sensitive to the degree of degradation: The more severe the acoustic degradation, the more enhanced alpha power in the centro-parietal region was observed (Obleser et al., 2012a). An increase in power in the alpha band has also been linked to attentional control mechanisms used by listeners to ignore irrelevant speech (Wöstmann et al., 2017) or stimuli (Jensen & Mazaheri, 2010), and with increased cognitive load in response to the processing of reduced word forms (Drijvers et al., 2016). Together these studies suggest that an increase in alpha power is indicative of increased cognitive load in the auditory domain and may reflect difficulties in semantic activation and retrieval. However, as pointed out by Van Engen and Peelle (2014), the cognitive effort caused by listening to accented speech remains largely unexplored, which will be one of the main focuses in the current study.

The present study

The main goal of the current study is to understand the underlying neurophysiological mechanisms of native- and nonnative-accented sentence processing in noisy conditions relative to quiet conditions. Specifically, using an auditory semantic violation paradigm, we tested young adults who listened to native- and nonnative-accented sentences with or without semantic anomalies, presented in quiet or in background noise (multi-speaker babble). To follow up on prior behavioral research in nonnative-accented sentences in noise processing (Bent, 2018; Bent & Atagi, 2015, 2017), we also measured behavioral comprehension accuracy by asking participants to verbally repeat back the sentence they just heard. The statistical analyses are divided into four parts: behavioral verbal repetition accuracy (measuring off-line comprehension), ERP analyses of the N400 component (associated with semantic access), alpha power increase (index of online working memory load), and exploratory analyses of theta power changes (associated with semantic processing).

Based on the literature reviewed above, we anticipated the following. For verbal repetition accuracy, we expected to replicate the behavioral findings of adult participants in Bent and Atagi (2015): Lower repetition accuracy for nonnative-accented sentences and for sentences presented in noise, and that accuracy would decline even more drastically when participants have to repeat nonnative-accented sentences presented in noise. With respect to ERP analyses of the N400, we predicted a delayed and smaller N400 effect for semantic violations in nonnative-accented sentences relative to native-accented sentences, in line with prior ERP research

(Gosselin et al., 2021; Grey & Van Hell, 2017; Schiller et al., 2020; Wambacq et al., 2023). Second, for the nonnative-accented speech in noise condition, in line with prior behavioral findings, we predicted a smaller or even complete absence of the N400 effect. Regarding analyses of alpha power, we predicted that both nonnative accent and background noise will elicit higher alpha power activity, indexing higher cognitive demand during online processing (albeit there is no prior research that has examined changes in neural oscillations associated with nonnative-accented speech processing). Please note that the semantic manipulation is irrelevant to the hypotheses regarding alpha power changes: we measure alpha power changes as an index of online listening effort, associated with nonnative speech processing and speech in noise processing. Lastly, for the more exploratory analyses on theta power changes we expected to observe a theta power increase when processing semantic anomalies in the 'ideal listening' condition (i.e., native in quiet). If background noise negatively impacts lexical-semantic access in a similar fashion as predicted for the N400 modulation, we expected to see a smaller theta power increase for semantic anomalies presented in noise than in quiet. Likewise, we expected a smaller theta power increase for semantic anomalies in nonnative- than in native-accented sentences.

Methods

Participants

36 Monolingual, right-handed young adults (24 female, 11 male, one non-binary; M_{Age} = 19.49, *SD* = 1.00), all undergraduate students from a large public American university, participated in the study. Participants were either recruited through the SONA Psychology Department participant pool or through word of mouth. The participation of one participant

(female) was discontinued because of equipment errors and thus excluded from data analyses, leading to a final sample size of 35.

In addition to self-identification as monolingual American-English speakers with limited prior exposure to nonnative accents during recruitment, all participants completed the Language History Questionnaire (LHQ3; Li et al., 2020). All participants reported their country of origin to be the United States, and half of them reported not knowing a second language (L2). For the 18 participants who reported knowing an L2, their L2 proficiency level was considered low: on average, they reported their ability in L2 listening, speaking, reading and writing to be 2.12, 2.18, 2.47, and 2.06 respectively (out of a 7-point Likert scale ranging from "1 - very poor" to "7 - excellent"). In terms of exposure to nonnative accents, all reported limited exposure to nonnative accents, and only two participants reported experience of living in foreign countries, both for a brief period of time.

Materials

Sentences. We adopted the semantic violation paradigm and thus there were two sentence types: half the sentences were semantically well-formed (e.g., "*Kevin* ***reached*** *into his* ***<u>pocket</u>*** *to get the* ***keys***") and the other half contained a semantic anomaly (e.g., "*Kevin* ***reached*** *into his* ***<u>funeral</u>*** *to get the* ***keys***"). A total of 288 unique sentences were created and counterbalanced across conditions of Accent (American, Chinese), Background (quiet, multi-talker babble noise) and Semantics (semantically well-formed, semantically anomalous), with 36 sentences in each of the eight conditions.

On average, these sentences contained 11.33 words (*SD* = 1.51). Target words in each sentence (underlined in the example sentences above; "pocket" and "funeral") were matched on word length and lexical frequency. Following prior accented speech in noise processing literature

(Bent, 2018; Bent & Atagi, 2015; Bradlow & Bent, 2008), to compute the behavioral measure of verbal repetition accuracy, three to five keywords (bolded words in the example sentences above) in each sentence were selected, including the target word; keyword identification accuracy was calculated and expressed as percentage out of all keywords.

Speaker selection and recordings. The auditory recording of the sentences took place in a sound-proof chamber using a Fostex DC-R302 recorder and a head-mounted Audix HT5 condenser microphone. The microphone was adjusted and positioned about 1.5 inches away from each speaker's mouth. The data were digitized at 44.1 kHz, 16-bit. All sentences were recorded by four female speakers (two with an American-English accent, and two with a Chinese-English accent) and normalized for intensity to 75 dB, in all conditions. Both American-English accented speakers were monolinguals who grew up speaking only American English. The two Chinese-accented speakers were native speakers of Mandarin Chinese, and spoke English with a notable Chinese accent. Each speaker recorded all semantically well-formed and semantically anomalous sentences (i.e., 576 sentences per speaker). All speakers were instructed to read aloud each sentence three times, using their natural speed, volume, and prosody.

The average duration of sentences recorded by American-English speakers was 3.02s, while those recorded by Chinese-accented speakers was 4.08s. This difference in duration is typical for speech produced by native- and nonnative-accented speakers, as slower speech rate is an inherent property of nonnative-accented speech (e.g., Baese-Berk & Morrill, 2015; Grey et al., 2019; Hanulíková et al., 2012; Romero-Rivas et al., 2015). To verify that sentence duration did not differ across other conditions, an Accent (native, nonnative) by Background (quiet, noise) by Semantics (well-formed, anomalous) repeated-measures ANOVA was conducted with sentence duration as the dependent variable. The only effect that was significant was the main effect of Accent ($F(1, 287) = 4550.00$, $p < .001$). No other main or interaction effects were significant (all ps > .35 or higher).

After the recording, the first author listened to the .WAV files of each speaker in Praat (Boersma & Weenink, 2023), and selected the most natural utterance for each sentence. Then, all individual sentences were segmented and extracted into single sound files and were normalized to 75dB. The sentence materials are made available on OSF.

Background noise. The multi-talker babble (MTB) noise used in the current study is an eight-talker English babble (total length 70.09s) that had also been used in Holt and Bent (2017). Half of the talkers were men, and the other half were women. Regarding talker accent, half of the talkers were native American-English speakers (2 male, 2 female) and the other half were native speakers of Mandarin Chinese (2 male, 2 female). Due to the large number of talkers (n = 8), the babble has become unintelligible, as evidence by studies comparing masking effects of different types of background noise (Freyman et al, 2005; Hoen et al., 2007; Rosen et al., 2013; see Mattys et al., 2012 for more information on informational masking).

An SNR level of +5 dBs was selected for the EEG sentence listening task, and sentence stimuli were embedded in a random segment of MTB noise, using scripts implemented in Praat (Boersma & Weenink, 2023). Following prior literature (e.g., Bent & Atagi, 2015, 2017; Holt & Bent, 2017), to mimic real-life noisy environments, 500 ms of MTB noise was added before and after the sentence. Correspondingly, 500 ms of a quiet period was also added before and after the sentences in the quiet condition.

Debriefing Survey. After completion of the EEG sentence listening task, participants were given a debriefing survey asking whether they detected a difference in the accent of the four speakers, following Grey and Van Hell (2017). If participants selected ‘no’, meaning that they did not detect a difference in accent, they were prompted to submit the survey. If participants selected 'yes', they were asked to identify the accent of each speaker ('indicate where the speaker comes from'). This portion of the survey was 'free response', meaning that participants could type any response with respect to the accent identity of the speaker(s). Additionally, participants rated the strength of each accent on a Likert scale where *1 = no accent* and *7 = very strong accent,* and how difficult it was to understand each speaker (*1 = very easy to understand* and *7 = very difficult to understand*).

Procedure

This study was approved by the Institutional Review Board of the university. Participants were tested in a single session, lasting approximately 3.5 hours. After providing informed consent, participants completed the Language History Questionnaire (LHQ3; Li et al., 2020), assessing their linguistic background and language use.

After completion of the LHQ3, the EEG sentence listening task was administered. Participants were seated in a comfortable chair in a sound-attenuated chamber, approximately three feet away from a computer monitor. An elastic cap with a total of 32 active electrodes was placed on the participant's head. Prior to the start of the experimental task, the experimenter read aloud the instructions and completed 12 practice trials with the participant, and answered any questions being raised.

Then, participants listened to a total of 288 sentences. Before the presentation of each sentence, participants saw a "Ready?" screen for them to blink or rest their eyes. Only when they were ready to proceed, they pressed the middle button on the button box, which triggered the auditory presentation of the next sentence. During sentence presentation, participants were instructed to look at the fixation cross ("+") in the center of the screen and were asked to relax and minimize eye-movements and blinks while listening to the sentences. After each sentence presentation, there was a 50% chance that participants saw the “Repeat” screen, during which they repeated aloud the sentence they just heard. There was no time limit for repeating back the sentence and the participants pressed the middle button after producing their verbal response (see Figure 1 for a schematic overview of each trial).

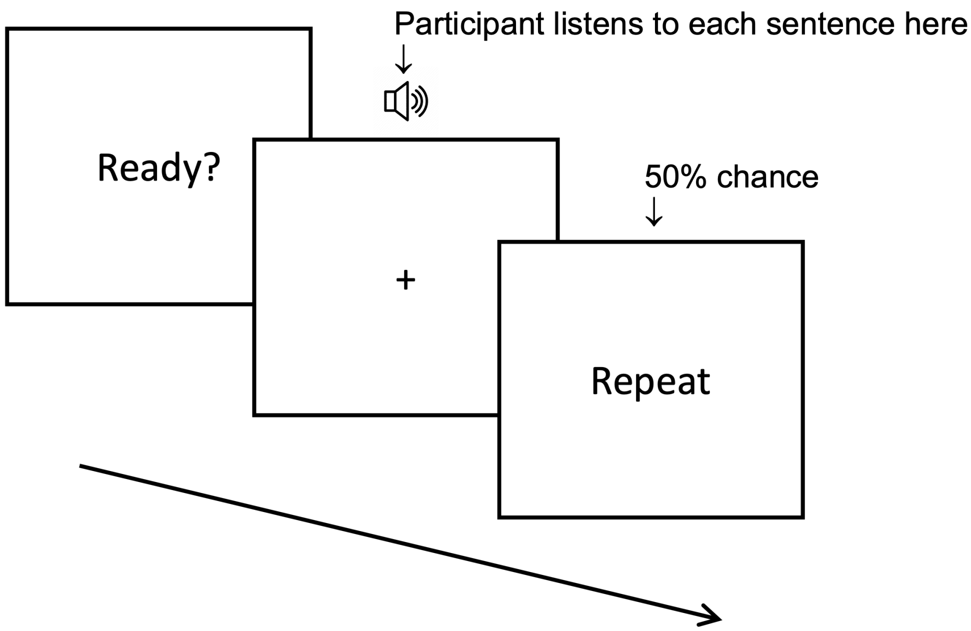

Figure 1. Schematic overview of the EEG Sentence Listening Task

Immediately following the EEG sentence listening task, the debriefing survey was administered (described in the ***Materials*** section), in which participants were asked to identify each accent and rate the accent strength and intelligibility of each of the four speakers.

EEG data acquisition and preprocessing

An elastic cap (BrainProducts ActiCap, Germany) with 32 active Ag/AgCl electrodes was placed on the participant’s head: four electrodes located along the midline (Fz, Cz, Pz, Oz) and 28 electrodes on the lateral sites (FP1/2, F7/8, F3/4, FC5/6, FC1/2, T7/8, C3/4, CP5/6,

CP1/2, P7/8, P3/4, PO9/10, O1/2, M1/2). An additional four eye electrodes were placed above and below the left eye, and at the outer canthus of each eye, to screen for ocular artifacts. Electrode impedances were kept below 5kΩ. EEG signals were amplified by a NeuroScan SynAmps RT amplifier using a .05-100Hz bandpass filter (first-order Butterworth with a 6 dB/octave roll-off) and digitized with a 500Hz sampling rate. Electrodes were online referenced to a vertex reference (ground electrode AFz, reference electrode FCz).

All pre-processing steps and analyses were performed using the EEGLAB (v2021.1; Delorme & Makeig, 2004) and ERPLAB (v8.30; Lopez-Calderon & Luck, 2014) toolboxes implemented in MATLAB R2022a (The MathWorks, Inc). Continuous EEG data was first band-pass filtered between 0.1 and 40 Hz using a 2nd-order noncausal IIR Butterworth digital filter (12 dB/octave roll-off). Unsystematic artifacts in continuous EEG data caused by muscle activity were detected and removed during manual inspection. Bad channels were identified via visual inspection and with the help of the Clean_Rawdata plugin (Kothe & Makeig, 2013; https://github.com/sccn/clean_rawdata), with minimum acceptable channel correlation at 0.8 and maximum acceptable line noise relative to its signal at four standard deviations. Continuous data were then re-referenced to the algebraic mean of activity over the left (M1) and right (M2) mastoids. After re-referencing, Independent Component Analysis (ICA) was performed and components containing prototypical ocular artifacts were removed from the data. Following ICA, missing channels were interpolated using the spherical spline method implemented in EEGLAB.

For ERP analyses, preprocessed EEG data were segmented into 1200 ms epochs (−200 to 1000 ms), time-locked to the onset of the semantically well-formed or anomalous target words. Epochs with activity exceeding ±100 µV at any electrode site were automatically flagged using a peak-to-peak moving window (window size: 200 ms; window step: 100 ms) procedure in

ERPLAB. All epochs were then manually inspected to ensure quality. The average epoch rejection rate was less than 5.23% (4.99%, 5.23%, 4.66%, 3.60%, 4,10%, 3.51%, 3.43%, 3.83% for each of the eight conditions). A one-way ANOVA with number of rejected epochs as the dependent measure yielded no effect of Condition (Greenhouse–Geisser corrected: $F(3.73, 122.97) = 1.19$, $p = .32$), indicating that rejection rates did not differ across conditions. ERPs were then averaged offline for each participant at each electrode site in each experimental condition, from 200 ms prior to stimulus onset (pre-stimulus baseline) to 1000 ms after stimulus onset.

To examine alpha power changes (8-12 Hz) as an index of listening effort, we conducted power spectral density (PSD) analyses. Power spectral density function (PSD) analysis provides the averaged strength of power as a function of frequency. We focused on alpha power changes in the EEG signal on the second word in the sentence, and well before the semantic anomaly/control target word was presented later in the sentence, to prevent that the semantic manipulation would confound any alpha power changes induced by accent or background noise. Thus, preprocessed EEG data were time-locked to the second word of each sentence and segmented into 1000 ms epochs (0 to 1000 ms), to obtain the average alpha power activity over the one-second period. The mean PSDs (expressed in $\mu V^2/Hz$) of all scalp electrodes were computed using the *spectopo* function of EEGLAB, which returns a PSD estimate via Welch's method.

To explore potential theta power changes (4-8 Hz) associated with hearing the semantically anomalous target word in the sentence, preprocessed EEG data was segmented into 3000 ms epochs (-1000 ms to 2000 ms), time-locked to the onset of semantically well-formed or anomalous target words. To detect event-related spectral perturbation (ERSP; Makeig, 1993), Morlet wavelet transformations were computed between 3 and 30 Hz, using the newtimef() function of EEGLAB, with linearly increasing number of cycles from three cycles for the lowest frequency (i.e., 3 Hz) and with a scale expansion factor of 0.8. A scaling factor of 0.8 indicates that the width of the wavelet used for the highest frequency (i.e., 30 Hz) is 0.2 (1 minus 0.8) of the width of the wavelet used at the lowest frequency, providing a good time-frequency trade-off for both low and high frequencies (Delorme & Makeig, 2004). The number of output times was set to 500 and the number of output frequencies was 200. Single trial normalization was performed, and power values were baseline corrected with respect to a 500 ms pre-stimulus onset.

Data analytic plan

For all analyses conducted, only significant effects ($p < .05$) are reported. Besides the planned a priori analyses, when a significant interaction effect emerged ($p < .05$), post-hoc ANOVA tests were conducted. Mauchly's sphericity tests was conducted when the variable has more than two levels, and Greenhouse–Geisser correction was used when sphericity was violated.

Additionally, to supplement null results, we computed Bayes factors (BF01), which provide a measure of the likelihood that the data are assuming the null hypothesis is true, relative to the alternative hypothesis (e.g., Schmalz et al., 2023). When BF01 > 1, this indicates more support for the null hypothesis by 'BF01' times. We used the BayesFactor R package (Morey & Rouder, 2015) which calculates a default Bayes factor with a wide Cauchy distribution (scale of effect = 0.707).

Verbal repetition analysis. The verbal repetition responses from the sentence listening task were recorded and coded offline by trained English-native undergraduate research assistants.

Number of keywords correctly identified were scored and percentages of correct keyword identification were calculated as the accuracy measure. Verbal repetition accuracy scores (in percentages) were submitted to an Accent (native, nonnative) by Background (quiet, noise) by Semantics (well-formed, anomalous) ANOVA, treating all variables as within-participant factors.

EEG data analyses: N400. In terms of scalp distribution, the N400 effect is typically most pronounced in central and parietal regions. We therefore conducted a Mass Univariate Analysis (Groppe et al., 2011) on the anomalous – well-formed difference wave across experimental conditions, only including centroparietal electrodes, following Jończyk et al. (2020). In this conservative mass univariate approach (recommended by Luck & Gaspelin, 2017), the ERPs from the critical condition are submitted to a repeated-measure two-tailed permutation test based on the tmax statistic (Blair & Karniski, 1993), using a family-wise alpha level of 0.05, to detect reliable mean difference of amplitudes in the time windows between 100 and 900 ms. The default setting was used, which conducted 2500 permutations. This procedure is characterized by a strong control of familywise error rate.

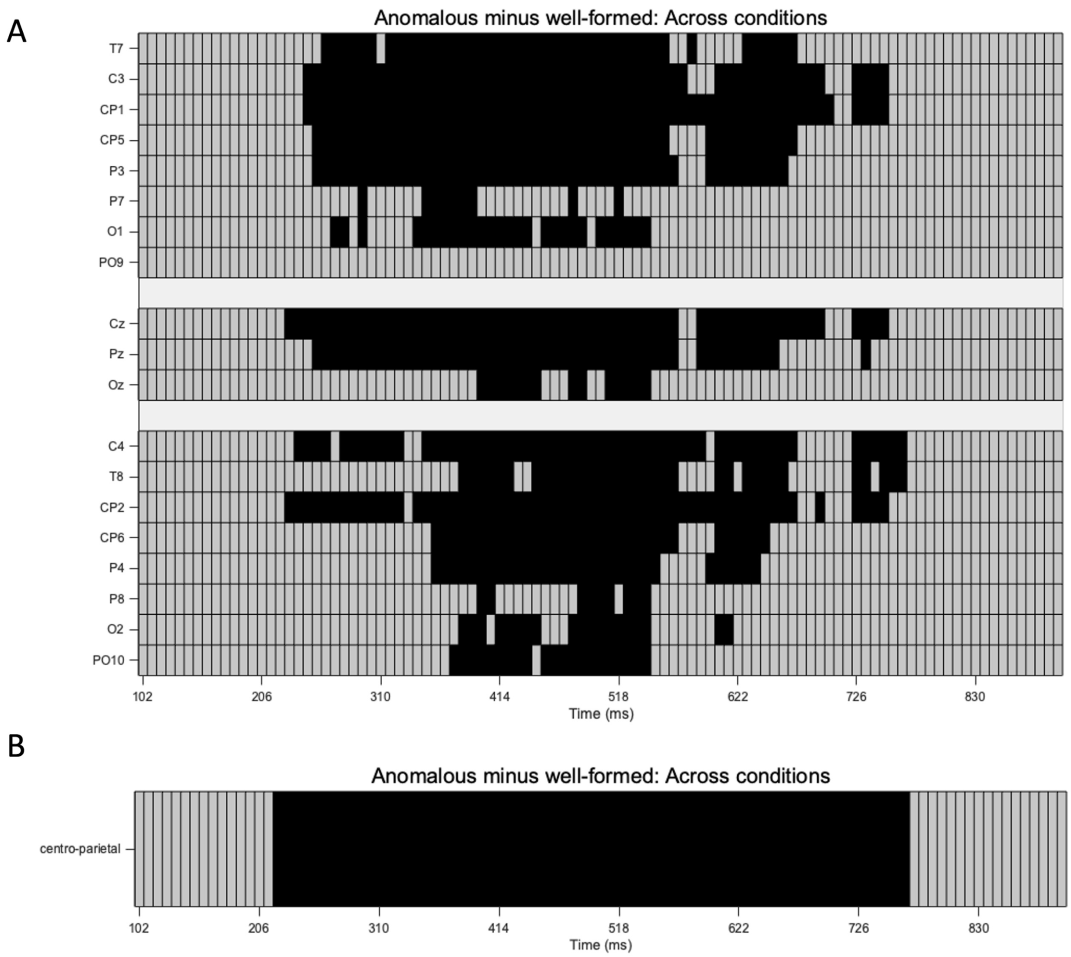

Figure 2. Results of the tmax permutation test from mass univariate analyses of the anomalous minus well-form comparison in centroparietal electrodes comparison of the ERP data, collapsed across experiment conditions. Significant t-tests for negative ERP differences are represented in black. No positive ERP differences were observed.

As a result of this procedure (see Figure 2 Panel A), 10 electrodes were selected and grouped into one centro-parietal ROI (C3, Cz, C4, CP5, CP1, CP2, CP6, P3, Pz, P4), where the effect was found to emerge within the 200 ms to 800 ms time window post-stimulus onset (see Figure 2 Panel B). Omnibus ANOVAs were then computed with Accent (native, nonnative), Background (quiet, noise), and Semantics (well-formed, anomalous) as within-participant factors, and with average amplitude in the 200-800 ms time-window as the dependent measure. More importantly, to test our a priori hypotheses, planned tests of the Semantics effect were conducted separately in each of the four conditions: native in quiet, nonnative in quiet, native in noise, and nonnative in noise.

Besides the magnitudes of the N400 effect, we also examined the onset of the semantic access and activation process (i.e., onset of the N400), individually for each of the four Accent by Background conditions. Specifically, prior ERP literature evidenced that the onset of the N400 was modulated by nonnative accent (e.g., Grey & Van Hell, 2017) and by background noise (e.g., Coulter et al., 2021). We therefore also planned tmax permutation tests on the 'anomalous – well-formed' difference wave in each of the four conditions: native in quiet, nonnative in quiet, native in noise, and nonnative in noise, in order to capture a robust emergence of the N400 effect; the outcome of these four tmax permutation tests are reported in the Results section.

EEG data analyses: alpha power (8-12 Hz). Enhanced alpha power activity associated with higher listening effort has been found to emerge primarily in parieto-occipital regions (e.g., Hunter, 2020; Jokisch & Jensen, 2007; Meyer et al., 2013; Obleser et al., 2012b; Obleser & Weisz, 2012). We therefore grouped 10 electrodes into one single parieto-occipital ROI (P7, P3, Pz, P4, P8, PO9, O1, Oz, O2, and PO10). To examine the effect of background noise and of nonnative accent on alpha power changes, omnibus ANOVAs were computed with Accent (native, nonnative) and Background (quiet, noise) as within-participant factors, and with alpha PSDs as the dependent measure.

EEG data analyses: theta power (4-8 Hz). As reviewed in the Introduction, prior studies on enhanced theta power associated with semantic violation focused on language processing during ideal reading (e.g., Davidson & Indefrey, 2007; Hagoort et al., 2004; Hald et al., 2006; Pu et al., 2020; Wang, Zhu, et al., 2012) or listening (Schneider & Maguire, 2019) conditions, without manipulations of accent or background noise. To capture how accent and background types modulate theta power during the semantic integration process of the target word, a more exploratory analytical approach was taken by inspecting the spectral topographies first in each condition to select corresponding ROIs and time-window of interest for analysis.

Results

Behavioral results

Accent detection, identification, and ratings (*N* = 35)

Only one out of the 35 participants reported not detecting any differences in accents. For the two American-accented English speakers, participants were highly accurate in identifying speaker accents. More specifically, 31 and 32 participants out of the 35 participants, respectively, accurately identified the accent to be American. On average, participants rated the accent strength of the two American-accented speakers to be low (M = 2.26 out of 7, SD = 1.86) and highly intelligible (M = 1.76 out of 7, SD = 1.90).

With regard to the two Chinese-accented English speakers, accent identification success was much lower: out of the 35 participants, only 5 and 11 participants, respectively, successfully identified the accent of each of the two speakers to be Chinese. Still, the majority of participants (23 and 24 out of 35, respectively, for the two speakers) identified the accent to be Asian or from other Asian countries. On average, participants rated the accent strength of the Chinese-accented speakers to be quite high (M = 5.07 out of 7, SD = 1.50) and fairly intelligible (M = 3.98 out of 7, SD = 1.64).

Verbal repetition accuracy (*N* = 34)

The verbal repetition recording sound file of one participant was corrupted due to technical errors and was excluded from analysis; the resulting sample size was 34. The ANOVA on the verbal repetition accuracy scores yielded main effects of Accent ($F(1, 33) = 494.60$, $p < .001$, $\eta_{partial}^2 = .94$, 90%CI [.90, .95]), Background ($F(1, 33) = 259.60$, $p < .001$, $\eta_{partial}^2 = .89$, 90%CI [.81, .92]), and Semantics ($F(1, 33) = 139.80$, $p < .001$, $\eta_{partial}^2 = .81$, 90%CI [.69, .86]). In general, participants were more accurate when repeating back native- (87%) versus nonnative-accented (67%) sentences, sentences presented in quiet (88%) versus noise (66%), and semantically well-formed (83%) versus anomalous (71%) sentences. Additionally, there was an Accent x Background x Semantics three-way interaction effect ($F(1, 33) = 11.84$, $p = .001$, $\eta_{partial}^2 = .26$, 90%CI [.07, .44]); follow-up tests were conducted for each of the background conditions, as reported next.

Table 1. Mean Accuracy Scores (and SDs) of Verbal Repetition

	Semantically well-formed		Semantically anomalous	
	Native	Nonnative	Native	Nonnative
Quiet	.97 (.04)	.89 (.06)	.94 (.07)	.74 (.11)
Noise	.85 (.09)	.61 (.16)	.72 (.13)	.46 (.14)

Quiet condition only. Follow-up tests on the quiet condition only revealed an Accent x Semantics interaction ($F(1, 33) = 56.93$, $p < .001$, $\eta_{partial}^2 = .63$, 90%CI [.44, .73]). As can be seen in Table 1 and Figure 3, the effect of semantics was stronger for nonnative-accented sentences ($F(1, 33) = 90.51$, $p < .001$, $\eta_{partial}^2 = .73$, 90%CI [.58, .80]) than for native-accented sentences ($F(1, 33) = 13.56$, $p < .001$, $\eta_{partial}^2 = .29$, 90%CI [.09, .46]). So in the quiet condition participants had significantly more difficulty repeating back anomalous sentences than well-formed sentences, and this difference was particularly pronounced for the nonnative-accented sentences.

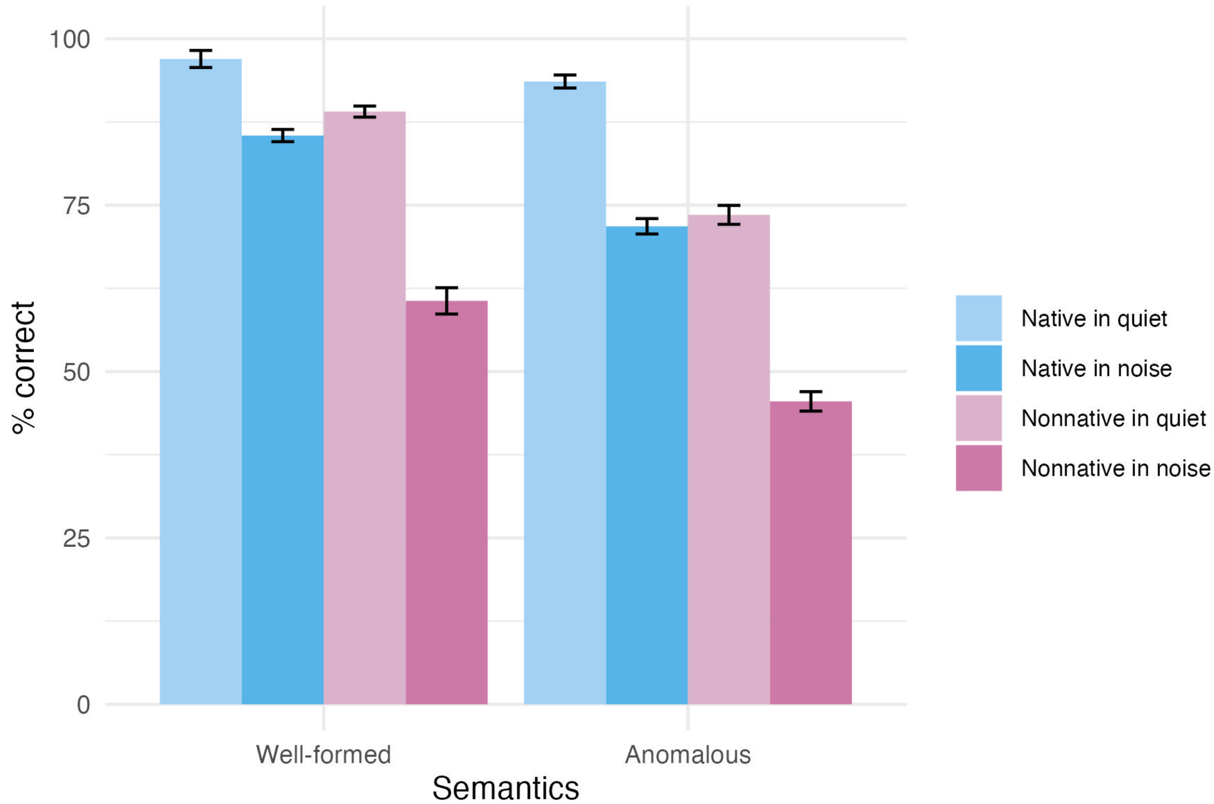

Figure 3. Verbal repetition accuracy scores. Error bars are ±SE.

Noise condition only. Only the main effects of Accent (F(1, 33) = 387.10, p < .001, $\eta_{partial}^2$ = .92, 90%CI [.87, .94]) and of Semantics (F(1, 33) = 118.4, p < .001, $\eta_{partial}^2$ = .78, 90%CI [.65, .84]) were significant; the Accent x Semantics interaction was not significant (F(1, 33) = 0.30, p = .59, $\eta_{partial}^2$ = .01, 90%CI [0, .11], BF01 = 1.06). When speech was embedded in multi-talker babble noise, participants were more accurate in repeating native-accented than nonnative-accented sentences. Additionally, participants were better at repeating semantically well-formed sentences than anomalous sentences (in both Accent conditions).

EEG results (N = 34)

As mentioned in the Methods section, the EEG data from one participant were excluded because of a high artifact rejection rate (> 30%). Thus, the reported EEG data are based on 34 participants.

N400: index of lexical semantic access

Based on the mass univariate analysis, a 2 (Accent: native, nonnative) × 2 (Background: quiet, noise) × 2 (Semantics: well-formed, anomalous) ANOVA was conducted, with the mean N400 amplitude in the 200 - 800 ms time-window as the dependent variable.

This omnibus ANOVA revealed a main effect of Semantics ($F(1, 33) = 51.43$, $p < .001$, $\eta_{partial}^2 = .61$, 90%CI [.41, .71]), a Background x Semantics interaction effect ($F(1, 33) = 6.22$, $p = .02$, $\eta_{partial}^2 = .16$, 90%CI [.02, .33]), and an Accent x Semantics interaction effect ($F(1, 33) = 7.05$, $p = .01$, $\eta_{partial}^2 = .18$, 90%CI [.02, .35]). We do not report further post-hoc tests of these interaction effects, since we have planned a priori tests in each of the four critical experimental conditions, as are reported in the next section.

N400 magnitudes in each accent by noise condition **(Figure 4).** Planned ANOVA tests of the Semantic effect were conducted in each of the four experimental conditions: native accent in quiet, native accent in noise, nonnative accent in quiet, and nonnative accent in noise.

Compared to semantically well-formed sentences, listening to semantic anomalies elicited a significant N400 effect in three conditions: native-accented speech in quiet ($F(1, 33) = 44.52$, $p < .001$, $\eta_{partial}^2 = .57$, 90%CI [.37, .69]), native-accented speech in noise ($F(1, 33) = 19.91$, $p < .001$, $\eta_{partial}^2 = .38$, 90%CI [.16, .53]), and nonnative-accented speech in quiet ($F(1, 33) = 15.66$, $p < .001$, $\eta_{partial}^2 = .32$, 90%CI [.11, .49]); see Figure 4A, left and right upper and left lower panels, respectively. In contrast, nonnative-accented anomalies embedded in noise did not elicit a significant N400 effect ($F(1, 33) = 2.06$, $p = .16$, $\eta_{partial}^2 = .06$, 90%CI [0, .21], BF01

= 1.59; see Figure 4A, lower right panel). This absence of the N400 effect indicates that semantic access is severed when listeners process nonnative-accented sentences embedded in multi-talker background noise. This aligns with the behavioral finding that verbal repetition accuracy was poorest in this experimental condition compared to the other three conditions.

To statistically compare the magnitude of the N400 effect across conditions, we computed the difference waves (anomalous – well-formed) for each of the four critical conditions: native in quiet, native in noise, nonnative in quiet, and nonnative in noise. The one-way ANOVA revealed a significant effect of Condition (Greenhouse–Geisser corrected: $F(2.8, 92.27) = 4.94$, $p = .004$, $\eta_{partial}^2 = .13$, 90%CI [.03, .22]). Post-hoc t-tests showed that the native in quiet, native in noise, and nonnative in quiet conditions elicited a larger N400 effect than the nonnative in noise condition (ts > 2.03, $p \leq .05$, ds > 0.35, BF01s < 0.86). Additionally, listening to native-accented sentences in quiet elicited a marginally larger N400 effect than the native accent in noise ($t(33) = 1.78$, $p = .08$, $d = 0.31$, 90%CI [.01, .59], BF01 = 1.31) and the nonnative accent in quiet ($t(33) = 1.96$, $p = .06$, $d = 0.37$, 90%CI [.07, .66], BF01 = 1.00) conditions. The N400 effect did not differ in magnitude between the native in noise and the nonnative in quiet conditions ($t(33) = 0.11$, $p = .91$, $d = 0.02$, 90%CI [-0.26, .30], BF01 = 5.41).

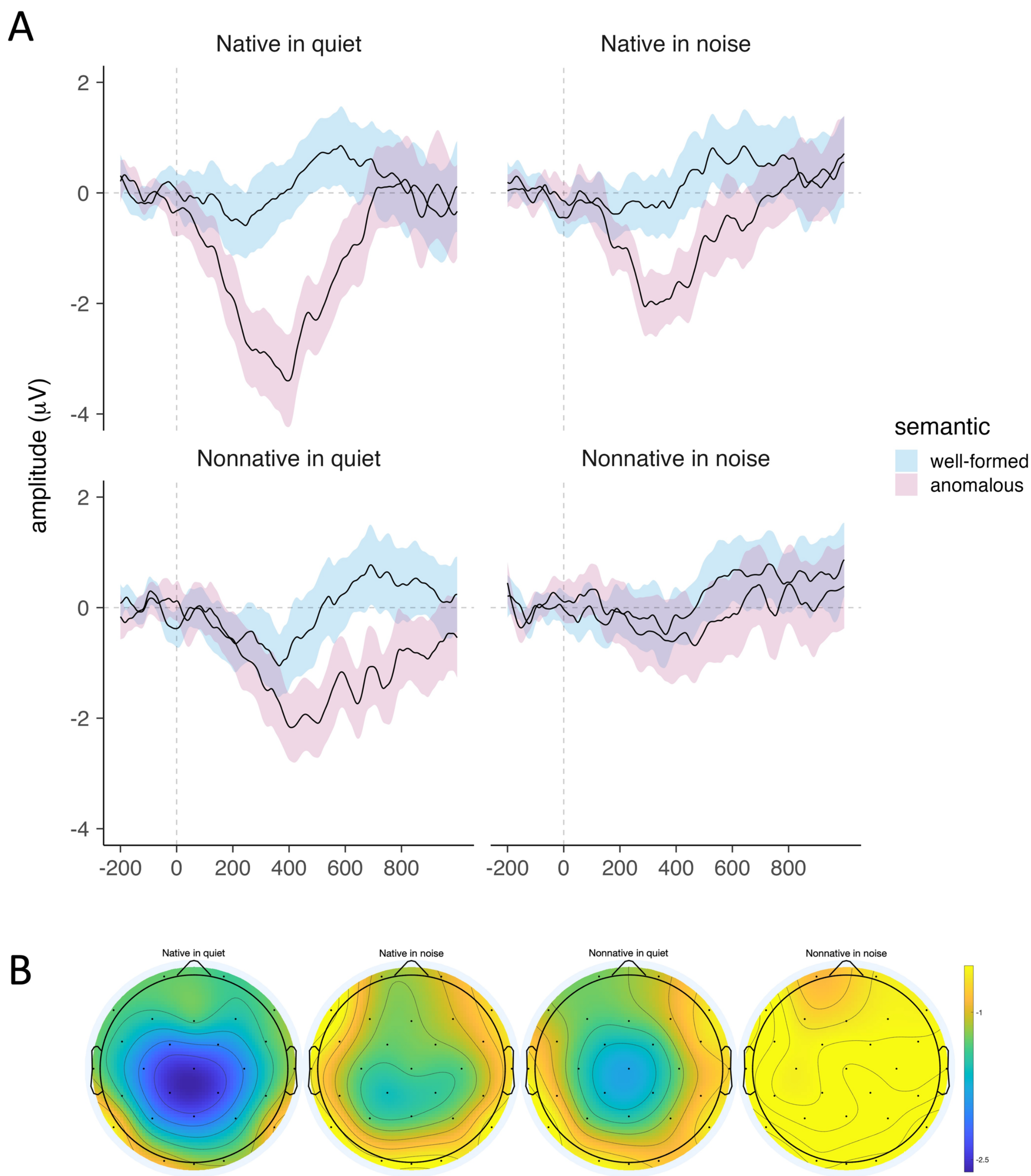

Figure 4. Panel A. N400 effects in the four critical accent by background conditions (upper left: native-accented sentences in quiet; upper right: native-accented sentences in noise; lower left: nonnative-accented sentences in quiet; lower right: nonnative-accented sentences in noise).

Responses were time-locked to target words that were either semantically well-formed or anomalous in the sentence context. Negative is plotted down. For presentation purposes only, waveforms were filtered with a 15 Hz low-pass filter. Shaded areas are ± SD. Panel B. Topographical distribution of the N400 effects in the four critical conditions.

N400 latency **(Figure 5).** To examine the onset of semantic access in each of the four conditions, planned tmax permutation tests were conducted in each of the four accent by noise conditions. These tests showed that the N400 time-window for the “ideal listening” condition (native accented speech in quiet) was 200 ms to 670 ms post stimulus onset, whereas nonnative-accented speech in quiet delayed the onset of the N400 effect (400 – 890 ms post stimulus onset). In comparison, the N400 time-window for native-accented speech in noise was slightly delayed and narrower (270 – 650 ms) compared to native-accented speech presented in quiet (200 – 670 ms). Taken together, this suggests that a nonnative accent casted a strong impact on the onset of the semantic access and activation process than background noise did. Lastly, the absence of a significant N400 amplitude effect in nonnative-accent sentence embedded in noise was paralleled in the tmax permutation test of this condition, in that no significant time window of an N400 effect was detected.

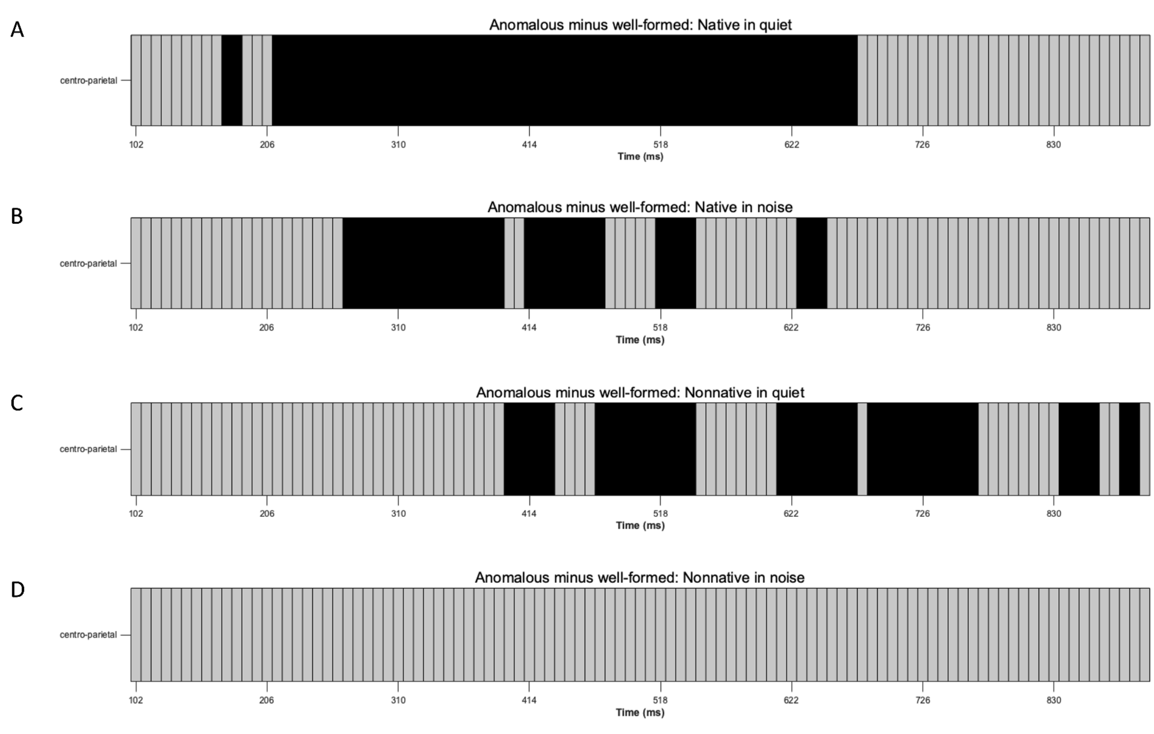

Figure 5. Results of tmax permutation tests from mass univariate analyses of the anomalous minus the well-formed waveforms, for each condition: native-accented sentences in quiet (Panel A), native-accented sentences in noise (Panel B), nonnative-accented speech in quiet (Panel C), and nonnative-accented sentences in noise (Panel D). Significant t-tests for negative ERP differences are represented in black. No positive-going ERP differences were observed.

Averaged alpha power increase: index of working memory load

The averaged alpha PSDs, measured from the onset of the second word of each sentence, were submitted to an Accent (native, nonnative) x Background (quiet, noise) within-participants ANOVA.

The main effect of Background ($F(1, 33) = 4.76$, $p = .04$, $\eta_{partial}^2 = .13$, 90%CI [.004, .30]) was significant: higher alpha power was associated with processing speech in multi-talker background noise ($M = 2.09$, $SD = 3.69$) than in a quiet environment ($M = 1.91$, $SD = 3.60$). We did not observe a main effect of Accent ($F(1, 33) < 0.001$, $p = .99$, $\eta_{partial}^2 < .001$, BF01 = 5.43) or an interaction between Background and Accent ($F(1, 33) = 0.001$, $p = .98$, $\eta_{partial}^2 < .001$, BF01 = 3.93), which we will elaborate on in the Discussion [1].

Event-related theta power changes: index of semantic integration

As outlined in the analytical plan, this part of the analysis is more exploratory in nature. To explore how accent and background types modulate theta power during semantic processing, we first computed the theta power difference value (anomalous minus well-formed) and inspected spectral topographies of power changes in each condition separately. Then, we statistically tested whether the Semantic effect on theta power changes was significant.

Native-accented speech in quiet (i.e., 'ideal listening' condition). As can be seen in Figure 6, theta power increased in the 550 ms to 900 ms time-window in the right centro-parietal

[1] The nonnative-accented speakers had slower speaking rates than the native-accented speakers (characteristic of nonnative-accented speech, see e.g., Baese-Berk & Morrill, 2015; Grey et al., 2019; Hanulíková et al., 2012; Romero-Rivas et al., 2015), so the same 1-second EEG epoch likely contains fewer phonemes in the nonnative- than in the native-accented speech condition. It could be argued that this may have driven the absence of an accent effect in the alpha power analysis. However, we did observe accent effects in the N400 analyses (for amplitude and latency) and in the theta power analyses, in which the EEG epoch is also (less than) 1-second. We therefore think that potential differences in the number of phonemes in the 1-sec epoch of the nonnative- and native-accented speech signals did not preclude the emergence of an accent effect in the alpha power analysis.

region (C4, CP1, CP2, CP6, P3, Pz, P4) when participants heard the semantic anomalous word in the sentence compared to the well-formed word; the effect of Semantics reached significance ($F(1, 33)= 11.80$, $p = .001$, $\eta_{partial}^2 =.26$, 90%CI [.07, .44]). Such enhanced theta synchronization when encountering semantic anomalies is in line with prior literature (e.g., Bastiaansen et al., 2008; Davidson & Indefrey, 2007; Hald et al., 2006; Pu et al., 2020), a point we will return to in the Discussion.

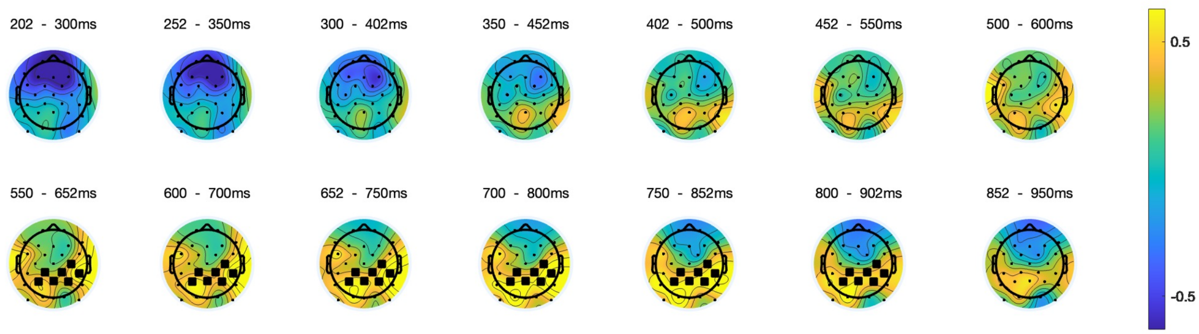

Figure 6. Topographies of theta power changes over time in the native in quiet condition (anomalous minus well-formed). Black squares represent electrodes selected for the ROI in the corresponding time window.

Native-accented speech in noise. Listening to native-accented sentences in background noise was also associated with increased theta power (see Figure 7), mainly in the right temporal region (FC6, C4, CP6, T8) at around 450 ms to 850 ms. The effect of Semantics reached significance ($F(1, 33) = 8.96$, $p = .005$, $\eta_{partial}^2 = .21$, 90%CI [.04, .39]), and emerged a bit earlier than in the quiet condition.

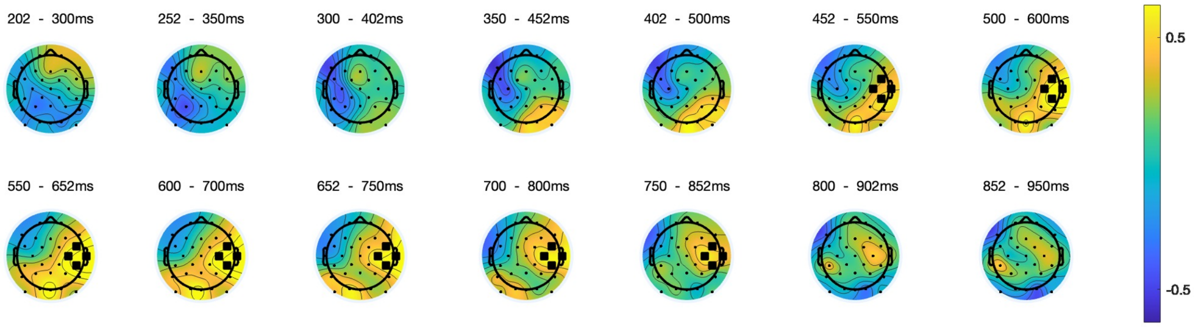

Figure 7. Topographies of theta power changes over time in the native in noise condition (anomalous minus well-formed). Black squares represent electrodes selected for the ROI in the corresponding time window.

Nonnative-accented speech in quiet. A decrease in theta power was observed for semantic anomalies relative to well-formed words in nonnative-accented sentences presented in quiet (see Figure 8; note that an increase in theta power was observed for anomalies in native-accented sentences). The effect emerged rather early at around 300 ms and ended at around 800 ms ($F(1, 33) = 5.18$, $p = .03$, $\eta_{partial}^2 = .14$, 90%CI [.01, .31]), in the right frontal region (Fz, F4, FC2, FC6).

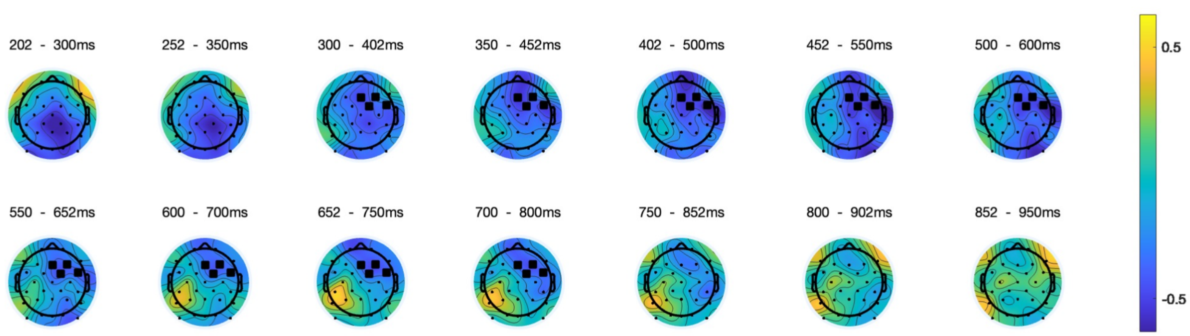

Figure 8. Topographies of theta power changes over time in the nonnative in quiet condition (anomalous minus well-formed). Black squares represent electrodes selected for the ROI in the corresponding time window.

Nonnative-accented speech in noise. Encountering semantic anomalies in nonnative-accented sentences embedded in noise (see Figure 9) was also associated with an early theta power decrease in the mid frontal region (F3, Fz, F4, FC1, FC2, C3, Cz, C4) in the 250 – 450 ms time-window, similar to the nonnative-accented speech in quiet condition. In the later 650 – 950 ms time-window, a theta power increase seemed to emerge in the left parietal region (P3, Pz, O1, Oz). However, the early theta decrease just reached statistical significance ($F(1, 33) = 4.22$, $p = .048$, $\eta_{partial}^2 = .11$, 90%CI [.001, .29]), while the later theta increase was only marginally significant ($F(1, 33) = 3.70$, $p = .06$, $\eta_{partial}^2 = .10$, 90%CI [0, .27], BF01 = 0.84).

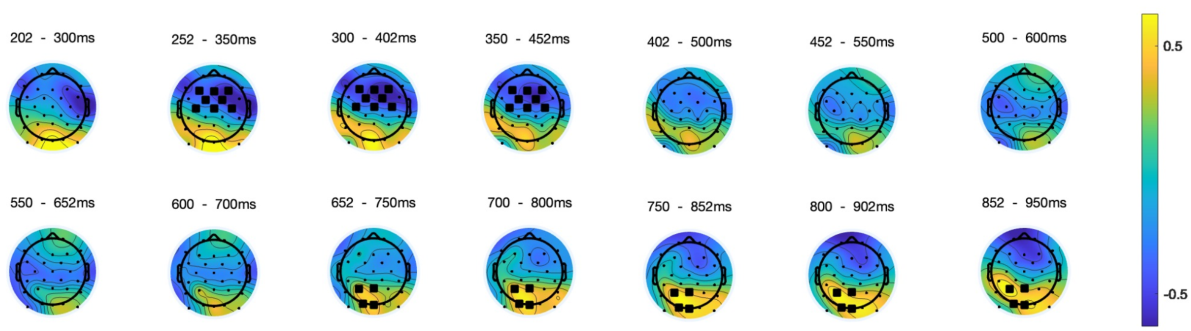

Figure 9. Topographies of theta power changes over time in the nonnative in noise condition (anomalous minus well-formed). Black squares represent electrodes selected for the ROI in the corresponding time window.

Discussion

In this study, we investigated the behavioral and neural correlates associated with the semantic processing of native- and nonnative-accented sentences embedded in noise (multi-talker babble) or presented in quiet. We also measured neural signatures of listening effort and cognitive demands when listeners process sentences in these four listening conditions. Behavioral measures of offline sentence comprehension (i.e., sentence repetition accuracy) showed that both listening to a nonnative accent and to speech embedded in multi-talker noise challenged listeners' sentence comprehension. Moreover, the adverse impact of multi-talker noise was exacerbated when listening to nonnative-accented sentences compared to native-accented sentences.

At the neural level, ERP analyses of the N400 effect indicated that both nonnative-accented speech and background noise impaired lexical-semantic access during online sentence processing, as indexed by the decrease in the N400 amplitude. The N400 onset latency analyses indicated that nonnative-accented speech delayed the semantic access and activation process, and more so than speech presented in multi-talker noise did. Furthermore, when listeners processed nonnative-accented speech in multi-talker noise, no semantic N400 effect was observed, in both the N400 amplitude and latency analyses. With regard to listening effort, we observed an enhanced alpha power activity (indexing increased working memory load) when participants listened to sentences embedded in a noisy background relative to presented in quiet. However, alpha power was not modulated by the accent manipulation. Lastly, we observed an increase in theta power activity when listeners encountered semantic anomalies when listening to native-accented sentences, both in quiet or in multi-talker noise, whereas semantic anomalies in nonnative-accented sentences were associated with a decrease in theta power activity, again both in quiet and in multi-talker noise.

Our behavioral verbal repetition findings paralleled earlier outcomes (Bent, 2018; Bent & Atagi, 2017): An unfamiliar nonnative accent and multi-talker noise impact speech comprehension, and this is further augmented when listeners were challenged with both. In fact, listeners were at ceiling in comprehending native-accented speech presented in quiet (> 90% accuracy), whereas comprehension accuracy decreased substantially when listeners were challenged with nonnative-accented speech embedded in multi-talker background noise (54% accuracy).

This decrease in comprehension accuracy after listeners heard the full sentence was associated with challenges in accessing the meaning of words while the sentence unfolds in real time, as evidenced by the ERP analyses. According to the TRACE model of speech perception (McClelland & Elman, 1986), listeners continually update lexical activation on the basis of the additional input. As the sentence unfolds, listeners build up expectations of the upcoming units based on the preceding context. A critical question in the current study is, as the discourse proceeds, whether speech produced with a nonnative accent or embedded in multi-talker noise achieves the same level of semantic access and activation as native speech presented in ideal listening conditions does. On the basis of the present N400 findings, the answer is likely to be 'No'.

To elaborate, in the current study, when participants listened to native-accented sentences in the quiet condition, semantically plausible lexical options become activated as the sentence unfolds in real time. Hearing an implausible item (i.e., an anomalous target word) elicited an enhanced N400 response, and this enhanced negativity extended into a later time window. When listeners were challenged with an unfamiliar nonnative accent or with multi-talker background noise, the semantic N400 effect became smaller, reflecting impaired lexical-semantic access and activation during online processing. Furthermore, in line with prior nonnative-accented speech N400 literature (e.g., Grey & Van Hell, 2017; Schiller et al., 2020; Wambacq et al., 2023), we found that semantic access was delayed for nonnative-accented compared to native-accented speech, as reflected in the N400 onset analysis. These combined N400 amplitude and onset findings indicate that listeners' ability to generate predictions about upcoming words based on semantic context is impeded when the acoustic features of the speech stream deviate from listeners' native/familiar phonetic representations, or when the speech is embedded in multi-talker background noise. It is also possible that listeners still engage in generating predictions but that the predictions have become less precise due to deviations and ambiguities in the speech signal (e.g., more potential lexical competitors have been pre-activated given the ambiguity of acoustic/phonological properties in the signal). Future studies could examine these interpretations more directly by manipulating sentential cloze probability.

This pattern of findings is in line with the speech comprehension model proposed by Van Engen and Peelle (2014): When there is a good acoustic match between the speaker and the listener, in our case American-English speaking participants listening to American-accented English sentences presented in quiet, the speech signal is processed relatively automatically with success. However, when the degree of acoustic match is reduced, for instance when speech is produced in an unfamiliar accent or when it is embedded in multi-talker background noise, listeners have to exert additional cognitive effort in order to compensate for the reduced acoustic match. When listeners encounter a nonnative accent embedded in multi-talker background noise, the acoustic mismatch becomes so strong that listeners do not have sufficient cognitive resources to fully compensate, leading to the severed online lexical-semantic access (demonstrated by the absence of a semantic N400 effect) and reduced offline sentence comprehension accuracy, as evidenced by listeners' faltering verbal repetition performance observed in the current study. Together these findings indicate that listeners' brains are less successful in generating timely and effective predictions based on the preceding semantic context as they listen to nonnative-accented sentences embedded in multi-talker noise that unfold in real-time.

Enhanced listening effort when the acoustic match becomes poor is also reflected in the observed changes in alpha power activity. Specifically, in the speech processing literature, increased alpha power has been associated with higher cognitive demands when listeners process degraded speech signals (Obleser et al., 2012b), when attending to a target speech stream while ignoring irrelevant speech (Wöstmann et al., 2017), and when processing speech in noisy backgrounds (Grant et al., 2022). In line with prior research and supporting our prediction, we indeed found that processing speech embedded in multi-talker noise is linked to an increased alpha activity in the parieto-occipital region. Notably, this effect was observed regardless of whether the speech was produced with a native or nonnative accent.

Listening to nonnative-accented sentences was thus not associated with increased alpha power activity; we present two possible accounts here. Possibly, this result is related to the experimental design of sentence presentation aiming to mimic naturalistic listening conditions. Specifically, when encountering a speaker in multi-talker background noise in real life, listeners are usually first immersed in the noisy backgrounds before the speaker starts talking. Behavioral studies have taken this into consideration in their experimental designs (e.g., Bent & Atagi, 2015) and exposed listeners to background noise first, before presenting them with either native- or nonnative-accented speech. In the current study, we followed a similar design by inserting a

500ms-long noisy or quiet period before the onset of each sentence utterance. Thus, before each sentence presentation, listeners were first exposed to the manipulated background type (multi-talker babble or quiet), after which they were presented with native- or nonnative accented speech. Even though the averaged alpha power level was measured for 1000 ms time-locked to the onset of the second word in the sentence, thus at a time the sentence unfolded either in quiet or in noise and was produced with either a native or a nonnative accent, the longer exposure to the specific background type may have nullified the accent manipulation.

Furthermore, background noise and nonnative accent are intrinsically different signals. Background noise is classified as a type of environmental speech degradation (Mattys et al., 2012), in which listeners experience signal loss due to the transmission process. Thus, listening to speech in noise might engage cognitive inhibition to keep the attention on the target speech while disregarding the background noise. In contrast, accented speech is conceptually different since variation in the acoustic signal does not result from an extrinsic source. Rather, the difficulty associated with listening to nonnative speech derives from online activation of phonetic and phonological competitors, instead of signal loss. As discussed in the Introduction, nonnative-accented speech posits a systematic mismatch between the incoming speech signal and listeners' phonological representations that have been built on the basis of native-accented speech input. More generally, these different sources help to understand the cognitive mechanisms underlying listeners' behavioral performance, in particular their challenges in repeating back nonnative-accented sentences immersed in multi-talker babble noise. Our findings indicate that comprehending speech embedded in multi-talker babble challenges cognitive resources ('working memory' and 'attention' in Van Engen and Peelle's (2014) model), whereas processing nonnative-accented speech presents challenges in mapping the deviating acoustic features in the nonnative-accented speech stream to L1-based phonological representations ('semantic integration' in Van Engen and Peelle's (2014) model).

Relatedly, Adank et al. (2015) reviewed and identified different neural networks supporting accented speech processing and speech in noise processing. They concluded that accented speech recruits additional brain areas outside the network activated for processing of speech in noise, time-compressed speech, and noise-vocoded speech. They also pointed out that networks in charge of accent processing might be confounded with the network that governs online cognitive resources (i.e., listening effort). They further argued for more controlled future investigations in which the specific contribution of accent processing can be teased apart from more general effortful processing. Taken together, it is possible that listening effort manifests itself in a different manner (and may be sensitive to different neural markers) for listening to nonnative-accented speech than for listening to speech embedded in multi-talker background noise, and more research is needed to further disentangle the underlying processing mechanisms associated with these different types of signals.

As outlined in the Introduction, semantic access during online sentence processing was measured with the frequently used N400 component, but we also explored power changes in the theta frequency band that have been associated with semantic processing in the literature. These analyses showed more pronounced differences for native- versus nonnative-accented speech than for quiet versus noisy conditions. An increase in theta power activity was associated with the processing of semantic anomalies in the native accent condition, both in quiet and in multi-talker babble. Such theta synchronization associated with semantic violations aligns with prior EEG literature (e.g., Davidson & Indefrey, 2007; Hagoort et al., 2004; Hald et al., 2006; Wang, Zhu, et al., 2012). Specifically, in the current study, the semantic effect associated with increased theta activity was detected in the right posterior region for the ideal listening condition (i.e., native in quiet condition), similar to what was found in Schneider and Maguire (2019) who also examined processing of semantic anomalies during listening (to speech produced by native speakers).

Listening to nonnative-accented sentences revealed a decrease in theta power activity in the mid-frontal region when listeners encountered semantic anomalies. To the best of our knowledge, the present study is the first to examine theta power changes during semantic processing of nonnative-accented speech, and we propose two possible interpretations for the observed theta power decrease. A first interpretation draws on the attention and cognitive control literature (for reviews, see Cavanagh & Frank, 2014; Nakamura-Palacios et al., 2023). Increased frontal midline theta rhythm (Fm-theta) has been frequently reported when individuals' task performance required mental concentration and high cognitive efforts, such as mental calculations (Asada et al., 1999; Ishii et al., 2014), the go/no-go task (Hong et al., 2020), an attention-demanding meditation procedure (Kubota et al., 2001), and working memory tasks (Kawasaki et al., 2010; Maurer et al., 2015). Fm-theta power changes during mental tasks have been localized in the anterior cingulate cortex and medial prefrontal cortex (e.g., Asada et al., 1999; Holroyd & Umemoto, 2016; Ishii et al., 2014; Töllner et al., 2017). Building on the link between increased/decreased Fm-theta and enhanced/reduced cognitive control, it is possible that when listeners experience challenges associated with listening to nonnative-accented speech, they decrease their efforts to keep their attention – which may be particularly true for listeners who are unfamiliar with nonnative-accented speech such as those tested in the present study. Thus, listeners may become less engaged in predictive processing as the sentence unfolds over time, which further tempers their attention to the semantic anomaly and reduces the efforts to access its meaning, as reflected in the decreased Fm-theta and the reduced N400. An alternative interpretation draws on studies that relate theta power decrease to the suppression of memory traces (Waldhauser et al., 2015) and to the suppression of potential alternative interpretations when processing grammatical violations (e.g., Rossi & Prystauka, 2020; Schneider et al., 2016). If acoustic/phonological ambiguities in the nonnative-accented speech signal impede listeners' ability to generate predictions about upcoming words based on the preceding sentence context (or if these predictions are less effective and activate a relatively high number of lexical competitors), the theta power decrease may reflect listeners' efforts to suppress lexical competitors of the critical target words. Note that, as we stated above, the analyses of theta power changes were more exploratory, and thus these interpretations need to be taken with caution and warrant further research.

Our findings have important real-world implications. Given that participants in this study were young college students at an R1 university who are at their peak of cognitive abilities, it is quite striking that listening to nonnative-accented speakers challenged access to lexical-semantic information as the sentence unfolded in real-time and their post-listening comprehension accuracy decreased (~ 53.5%). In comparison, when they listened to nonnative-accented sentences in quiet or to native-accented sentences in background noise, they were able to access lexical-semantic information in real-time and demonstrated decent levels of comprehension success (~ 80%). Our findings thus argue for the importance of auditory perceptual training, which has been found to improve speech comprehension (e.g., Loebach et al., 2009; Whitton et al., 2017), and further highlight the importance of adding multi-speaker babble noise to the training procedure. For instance, Zhang et al. (2021) randomly assigned participants to a training or control group, and the training group was directly trained on listening to words embedded in background babble noise. It was found that the trained participants demonstrated significant preto-post improvement at identifying words embedded in babble noise, as well as in car noise and in rain noise. Such improvement even maintained weeks after the training session, for words embedded in babble noise, but not in car or rain noise. In terms of adaptation to novel accents, prior studies have also demonstrated that short-term training in the laboratory and exposing listeners to nonnative-accented speech might alleviate the associated processing costs, and achieve talker-independent adaptation (see Bent & Baese-Berk, 2021 for a detailed review).

Additionally, for shared spaces like restaurants or medical settings, engineers and designers should be aware of the linguistic diversity among the occupants when designing acoustic features for the space. For instance, Lebo et al. (1994) reported that noise level of restaurants in San Francisco, CA, ranged from 59 to 80 dBA sound-pressure levels, with a mean loudness level of 71 dBA. Similarly, noise level of restaurants in Hong Kong ranged from 66.7 to 82.6 dBA, with a mean value of 73.9 dBA (To & Chung, 2015). In the hospital setting, noise levels were reported to be in the 50 to 60 dB(A) range (Busch-Vishniac et al., 2005), and the emergency department tends to exhibit noise levels that are even 5 to 10 dB(A) higher (Orellana et al., 2007). Given that the average intensity of conversational speech is typically around 65 dB, more stringent acoustic requirements might be necessary for listeners to achieve adequate speech comprehension performance, particularly when interacting with nonnative-accented speakers.

This study addressed novel questions, and yielded novel insights, regarding the neurocognitive correlates underlying the processing of nonnative-accented and native-accented sentences embedded in multi-talker background noise and in quiet conditions, integrating electrophysiological techniques (Event Related Potentials and time-frequency analyses) and behavioral comprehension accuracy measures. The present work also leaves certain questions unanswered that could be addressed in future research. First, the current participant sample was kept as homogeneous as possible with respect to limited prior experience with nonnative-accented speech, and we did not examine individual variability in prior experience with nonnative accented speech. Therefore, our outcomes may not generalize to listeners with extensive experience listening to nonnative-accented speech, including listeners living in more metropolitan areas. Future research may want to explore the generalizability of the present findings by testing listeners who have more experience with listening to nonnative-accented speech or who have had more exposure to this specific variety of nonnative-accented speech. Additionally, there might be individual differences in terms of preference over quiet or noisy conditions: Some individuals might prefer an absolutely quiet place in order to focus, while others find the presence of background noise helpful for concentration (e.g., using a white noise machine). Under these circumstances, it is possible that those who prefer and are used to having constant noise in the background are less affected by noisy conditions when processing speech, including nonnative-accented speech.

Furthermore, future research may seek to investigate the effect of nonnative accents and multi-talker background noise on second language (L2) processing in bilingual listeners. Specifically, the comprehension of nonnative-accented speech is arguably even more difficult for listeners when processing sentences in their L2. The available studies indeed found that L2 nonnative-accented speech processing is more difficult, and constrained by different factors, than L1 nonnative-accented speech processing (Grey et al., 2019; Lev-Ari et al., 2017; Weber et al., 2014). With one exception: the general difficulty of L2 speech comprehension has been found to be attenuated in listeners listening to their own accent (e.g., a Chinese native speaker listening to English sentences produced by another Chinese speaker), as compared to listeners who listen to L2 speech produced in a nonnative-accent different from their own accent, a phenomenon termed the Interlanguage Speech Intelligibility Benefit (ISIB: Bradlow & Bent, 2008; Hayes-Harb et al., 2008; Imai et al., 2005; see Gosselin et al., 2022 for EEG evidence). Recently, EEG studies have started to investigate how bilingual listeners process speech in their L2 embedded in noise (but produced by native-accented speakers; e.g., Coulter et al., 2021; Grant et al., 2022) and how bilingual listeners process nonnative- and native-accented L2 speech (presented in quiet conditions; Grey et al., 2019). It will be thus intriguing to combine these two lined of research and examine whether the ISIB attenuates, or can even serve as a protective factor against, undesirable effects of multi-talker background noise.

Lastly, we acknowledge some limitations in the design of the current study. First, due to practical consideration of experimental implementation, we were limited in the number of SNR levels we could have investigated, and thus only implemented one SNR level (+ 5 dBs) to compare to the quiet condition. Future research could directly examine these effects across a wide range of SNRs. Second, regarding the listening effort measure, we did not incorporate a direct behavioral measure of effort, and only measured the alpha activity, which is an implicit measure. Future research could include direct behavioral measures, such as a dual-task paradigm or a self-report measure (e.g., NASA Task Load Index; Hart & Staveland, 1988), to assess the relationship between neural indices and behavioral measures of listening effort. Third, EEG is known for its advantage in high temporal resolution, while being limited in spatial resolution. Researchers interested in spatial localization of these effects could endorse other neuroimaging techniques, such as fMRI or MEG, or use them in combination with EEG.

Taken together, in the current study, we have taken a novel approach by designing and implementing an EEG native- and nonnative-accented sentence listening task embedded in multi-talker babble noise, building on and extending the behavioral paradigm used in prior studies (e.g., Bent & Atagi, 2015), to measure both neural correlates during online sentence processing and behavioral offline sentence comprehension accuracy. We observed modulations of the neural signatures associated with listening to a nonnative accent and/or multi-talker background noise that were largely in line with our predictions, while also replicating earlier behavioral findings of verbal repetition accuracy. The current study also extended prior EEG investigations of accented speech, as in these studies participants were instructed to either listen passively (e.g., Romero-Rivas et al., 2015, 2016) or answer simple "yes/no" comprehension questions (e.g., Caffarra & Martin, 2019; Hanulíková et al., 2012) to keep their attention on the task. The novel approaches in the current study further bridge the gap between behavioral and electrophysiological studies on the comprehension of accented speech and speech in noise, and provide further insight into the relation between online neural processes and behavioral outcomes. Taken together, we found that listening to nonnative-accented sentences and speech embedded in multi-talker background noise is associated with challenges in semantic access of upcoming words as the sentence unfolds in real time and imposes cognitive efforts. Our findings indicate that the underlying neurocognitive mechanism might be different across accented speech processing and speech in noise processing: Comprehending speech embedded in multi-talker babble challenges cognitive resources, and processing nonnative-accented speech presents challenges in mapping the deviating acoustic features in nonnative-accented sentences to L1-based phonological representations.

References

Adank, P., Nuttall, H. E., Banks, B., & Kennedy-Higgins, D. (2015). Neural bases of accented speech perception. *Frontiers in Human Neuroscience*, *9*. https://doi.org/10.3389/fnhum.2015.00558

Asada, H., Fukuda, Y., Tsunoda, S., Yamaguchi, M., & Tonoike, M. (1999). Frontal midline theta rhythms reflect alternative activation of prefrontal cortex and anterior cingulate cortex in humans. *Neuroscience Letters*, *274*(1), 29–32. https://doi.org/10.1016/S0304-3940(99)00679-5

Baese-Berk, M. M., & Morrill, T. H. (2015). Speaking rate consistency in native and non-native speakers of English. *The Journal of the Acoustical Society of America*, *138*(3), EL223–EL228. https://doi.org/10.1121/1.4929622

Bastiaansen, M. C., Oostenveld, R., Jensen, O., & Hagoort, P. (2008). I see what you mean: theta power increases are involved in the retrieval of lexical semantic information. Brain and language, 106(1), 15-28. https://doi.org/10.1016/j.bandl.2007.10.006

Behrman, A., & Akhund, A. (2013). The influence of semantic context on the perception of Spanish-accented American English. *Journal of Speech, Language, and Hearing Research*, *56*(5), 1567–1578. https://doi.org/10.1044/1092-4388(2013/12-0192).

Bent, T. (2014). Children's perception of foreign-accented words. *Journal of Child Language*, *41*(06), 1334–1355. https://doi.org/10.1017/S0305000913000457

Bent, T. (2018). Development of unfamiliar accent comprehension continues through adolescence. *Journal of Child Language*, *45*(6), 1400–1411. https://doi.org/10.1017/S0305000918000053

Bent, T., & Atagi, E. (2015). Children's perception of nonnative-accented sentences in noise and quiet. *The Journal of the Acoustical Society of America*, *138*(6), 3985–3993. https://doi.org/10.1121/1.4938228

Bent, T., & Atagi, E. (2017). Perception of nonnative-accented sentences by 5- to 8-year-olds and adults: The role of phonological processing skills. *Language and Speech*, *60*(1), 110–122. https://doi.org/10.1177/0023830916645374

Bent, T., & Baese-Berk, M. (2021). Perceptual learning of accented speech. In J. S. Pardo, L. C. Nygaard, R. E. Remez, & D. B. Pisoni (Eds.), *The handbook of speech perception* (2nd ed., pp. 429-453). Wiley. https://doi.org/10.1002/9781119184096.ch16

Bent, T., & Holt, R. (2013). The influence of talker and foreign-accent variability on spoken word identification. *The Journal of the Acoustical Society of America*, *133*(3), 1677–1686. https://doi.org/10.1121/1.4776212

Boersma, Paul & Weenink, David (2018). Praat: doing phonetics by computer [Computer program]. Version 6.0.37 from http://www.praat.org/

Binder, J. R. (2016). Phoneme perception. In G. Hickok & S. L. Small (Eds), *Neurobiology of Language* (pp. 447–461). Elsevier. https://doi.org/10.1016/B978-0-12-407794-2.00037-7

Blair, R. C., & Karniski, W. (1993). An alternative method for significance testing of waveform difference potentials. *Psychophysiology*, *30*(5), 518–524. https://doi.org/10.1111/j.1469-8986.1993.tb02075.x

Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition*, *106*(2), 707–729. https://doi.org/10.1016/j.cognition.2007.04.005

Busch-Vishniac, I. J., West, J. E., Barnhill, C., Hunter, T., Orellana, D., & Chivukula, R. (2005). Noise levels in Johns Hopkins hospital. *The Journal of the Acoustical Society of America, 118*(6), 3629-3645. https://doi.org/10.1121/1.2118327

Caffarra, S., Gosselin, L., Thomas, T., & Martin, C. D. (2023). The neurocognition of foreign accent perception. In K. Morgan-Short & J. G. van Hell (Eds.), *The Routledge Handbook of Second Language Acquisition and Neurocognition* (pp. 397-411). Routledge. https://doi.org/10.4324/9781003190912

Caffarra, S., & Martin, C. D. (2019). Not all errors are the same: ERP sensitivity to error typicality in foreign accented speech perception. *Cortex*, *116*, 308–320. https://doi.org/10.1016/j.cortex.2018.03.007

Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, *18*(8), 414–421. https://doi.org/10.1016/j.tics.2014.04.012

Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *The Journal of the Acoustical Society of America*, *116*(6), 3647–3658. https://doi.org/10.1121/1.1815131

Coulter, K., Gilbert, A. C., Kousaie, S., Baum, S., Gracco, V. L., Klein, D., Titone, D., & Phillips, N. A. (2021). Bilinguals benefit from semantic context while perceiving speech in noise in both of their languages: Electrophysiological evidence from the N400 ERP. *Bilingualism: Language and Cognition*, *24*(2), 344–357. https://doi.org/10.1017/S1366728920000516

Davidson, D. J., & Indefrey, P. (2007). An inverse relation between event-related and time–frequency violation responses in sentence processing. *Brain Research*, *1158*, 81–92. https://doi.org/10.1016/j.brainres.2007.04.082

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Drijvers, L., Mulder, K., & Ernestus, M. (2016). Alpha and gamma band oscillations index differential processing of acoustically reduced and full forms. *Brain and Language*, *153–154*, 27–37. https://doi.org/10.1016/j.bandl.2016.01.003

Eberhard, D. M., G. F. Simons, & Charles D. Fennig (eds.). 2023. Ethnologue: Languages of the world. Twenty-sixth edition. Dallas, Texas: SIL International. Online version: http://www.ethnologue.com.

Federmeier, K. D. (2022). Connecting and considering: Electrophysiology provides insights into comprehension. Psychophysiology, 59(1), e13940. https://doi.org/10.1111/psyp.13940

Fernandez, C. B., Litcofsky, K. A., & van Hell, J. G. (2019). Neural correlates of intra-sentential code-switching in the auditory modality. *Journal of Neurolinguistics*, *51*, 17–41. https://doi.org/10.1016/j.jneuroling.2018.10.004

Fiedler, L., Seifi Ala, T., Graversen, C., Alickovic, E., Lunner, T., & Wendt, D. (2021). Hearing aid noise reduction lowers the sustained listening effort during continuous speech in noise—A combined pupillometry and EEG study. *Ear & Hearing*, *42*(6), 1590–1601. https://doi.org/10.1097/AUD.0000000000001050

Flege, J. E., & Bohn, O. (2021). The revised speech learning model (SLM-r). In R. Wayland (Ed.), Second language speech learning: *Theoretical and empirical progress* (pp. 3–83). Cambridge University Press.

Flege, J. E., MacKay, I. R. A., & Meador, D. (1999). Native Italian speakers' perception and production of English vowels. *The Journal of the Acoustical Society of America*, *106*(5), 2973–2987. https://doi.org/10.1121/1.428116

Flege, J. E., Schirru, C., & MacKay, I. R. A. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication*, *40*(4), 467–491.

https://doi.org/10.1016/S0167-6393(02)00128-0

Foxe, J. J., & Snyder, A. C. (2011). The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00154

Freyman, R. L., Balakrishnan, U., & Helfer, K. S. (2004). Effect of number of masking talkers and auditory priming on informational masking in speech recognition. *The Journal of the Acoustical Society of America, 115(5)*, 2246-2256. https://doi.org/10.1121/1.1689343

Gordon-Salant, S., Yeni-Komshian, G. H., & Fitzgibbons, P. J. (2010). Recognition of accented English in quiet and noise by younger and older listeners. *The Journal of the Acoustical Society of America*, *128*(5), 3152–3160. https://doi.org/10.1121/1.3495940

Gordon-Salant, S., Yeni-Komshian, G. H., Fitzgibbons, P. J., Cohen, J. I., & Waldroup, C. (2013). Recognition of accented and unaccented speech in different maskers by younger and older listeners. *The Journal of the Acoustical Society of America*, *134*(1), 618–627. https://doi.org/10.1121/1.4807817

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

Gosselin, L., Martin, C. D., González Martín, A., & Caffarra, S. (2022). When a nonnative accent lets you spot all the errors: Examining the syntactic interlanguage benefit. *Journal of Cognitive Neuroscience, 34*(9), 1650-1669. https://doi.org/10.1162/jocn_a_01886

Gosselin, L., Martin, C. D., Navarra-Barindelli, E., & Caffarra, S. (2021). The presence of a foreign accent introduces lexical integration difficulties during late semantic processing. *Language, Cognition and Neuroscience*, *36*(9), 1086–1106.

https://doi.org/10.1080/23273798.2021.1909084

Grant, A. M., Kousaie, S., Coulter, K., Gilbert, A. C., Baum, S. R., Gracco, V., Titone, D., Klein, D., & Phillips, N. A. (2022). Age of acquisition modulates alpha power during bilingual speech comprehension in noise. *Frontiers in Psychology*, *13*, 865857. https://doi.org/10.3389/fpsyg.2022.865857

Grey, S., Schubel, L. C., Mcqueen, J. M., & van Hell, J. G. (2019). Processing foreign-accented speech in a second language: Evidence from ERPs during sentence comprehension in bilinguals. *Bilingualism: Language and Cognition*, *22*(5), 912–929. https://doi.org/10.1017/S1366728918000937

Grey, S., & van Hell, J. G. (2017). Foreign-accented speaker identity affects neural correlates of language comprehension. *Journal of Neurolinguistics*, *42*, 93–108. https://doi.org/10.1016/j.jneuroling.2016.12.001

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review: Mass univariate analysis of ERPs/ERFs I: Review. *Psychophysiology*, *48*(12), 1711–1725. https://doi.org/10.1111/j.1469-8986.2011.01273.x

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition*, *1*(2), 131–149. https://doi.org/10.1017/S136672899800025X

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, *304*(5669), 438–441. https://doi.org/10.1126/science.1095455

Hald, L. A., Bastiaansen, M. C. M., & Hagoort, P. (2006). EEG theta and gamma responses to semantic violations in online sentence processing. *Brain and Language*, *96*(1), 90–105.

https://doi.org/10.1016/j.bandl.2005.06.007

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), 10. https://doi.org/doi.org/10.1162/jocn_a_00103

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human Mental Workload* (Vol. 52, pp. 139-183). North-Holland. https://doi.org/10.1016/S0166-4115(08)62386-9

Hayes-Harb, R., Smith, B. L., Bent, T., & Bradlow, A. R. (2008). The interlanguage speech intelligibility benefit for native speakers of Mandarin: Production and perception of English word-final voicing contrasts. *Journal of Phonetics*, *36*(4), 664–679. https://doi.org/10.1016/j.wocn.2008.04.002

Hoen, M., Meunier, F., Grataloup, C. L., Pellegrino, F., Grimault, N., Perrin, F., ... & Collet, L. (2007). Phonetic and lexical interferences in informational masking during speech-in-speech comprehension. *Speech Communication, 49(12),* 905-916. https://doi.org/10.1016/j.specom.2007.05.008

Holroyd, C. B., & Umemoto, A. (2016). The research domain criteria framework: The case for anterior cingulate cortex. *Neuroscience & Biobehavioral Reviews*, *71*, 418–443. https://doi.org/10.1016/j.neubiorev.2016.09.021

Holt, R. F., & Bent, T. (2017). Children's use of semantic context in perception of foreign-accented speech. *Journal of Speech, Language, and Hearing Research*, *60*(1), 223–230. https://doi.org/10.1044/2016_JSLHR-H-16-0014

Holt, R., Kung, C., & Demuth, K. (2018). Listener characteristics modulate the semantic processing of native vs. foreign-accented speech. *PLOS ONE*, *13*(12), e0207452. https://doi.org/10.1371/journal.pone.0207452

Hong, X., Sun, J., Wang, J., Li, C., & Tong, S. (2020). Attention-related modulation of frontal midline theta oscillations in cingulate cortex during a spatial cueing Go/NoGo task. *International Journal of Psychophysiology*, *148*, 1–12. https://doi.org/10.1016/j.ijpsycho.2019.11.011

Huang, B., & Evanini, K. (2016). Think, sink, and beyond: Phonetic variants and factors contributing to English *th* pronunciation among Chinese speakers. *Journal of Second Language Pronunciation*, *2*(2), 253–275. https://doi.org/10.1075/jslp.2.2.06hua

Hunter, C. R. (2020). Tracking cognitive spare capacity during speech perception with EEG/ERP: Effects of cognitive load and sentence predictability. *Ear & Hearing*, *41*(5), 1144–1157. https://doi.org/10.1097/AUD.0000000000000856

Imai, S., Walley, A. C., & Flege, J. E. (2005). Lexical frequency and neighborhood density effects on the recognition of native and Spanish-accented words by native English and Spanish listeners. *The Journal of the Acoustical Society of America*, *117*(2), 896–907. https://doi.org/10.1121/1.1823291

Ishii, R., Canuet, L., Ishihara, T., Aoki, Y., Ikeda, S., Hata, M., Katsimichas, T., Gunji, A., Takahashi, H., Nakahachi, T., Iwase, M., & Takeda, M. (2014). Frontal midline theta rhythm and gamma power changes during focused attention on mental calculation: An MEG beamformer analysis. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00406

Jensen, O., Gelfand, J., Kounios, J., & Lisman, J. E. (2002). Oscillations in the Alpha Band (9-12

Hz) Increase with memory load during retention in a short-term memory task. *Cerebral Cortex*, *12*(8), 877–882. https://doi.org/10.1093/cercor/12.8.877

Jensen, O., & Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Frontiers in Human Neuroscience*, *4*. https://doi.org/10.3389/fnhum.2010.00186

Jokisch, D., & Jensen, O. (2007). Modulation of gamma and alpha activity during a working memory task engaging the dorsal or ventral stream. *The Journal of Neuroscience*, *27*(12), 3244–3251. https://doi.org/10.1523/JNEUROSCI.5399-06.2007

Jończyk, R., Kremer, G. E., Siddique, Z., & van Hell, J. G. (2020). Engineering creativity: Prior experience modulates electrophysiological responses to novel metaphors. Psychophysiology, 57(10), e13630. https://doi.org/10.1111/psyp.13630

Kawasaki, M., Kitajo, K., & Yamaguchi, Y. (2010). Dynamic links between theta executive functions and alpha storage buffers in auditory and visual working memory: EEG oscillation for working memory system. *European Journal of Neuroscience*, no-no. https://doi.org/10.1111/j.1460-9568.2010.07217.x

Kothe, C. A., & Makeig, S. (2013). BCILAB: A platform for brain–computer interface development. *Journal of Neural Engineering*, *10*(5), 056014. https://doi.org/10.1088/1741-2560/10/5/056014

Kubota, Y., Sato, W., Toichi, M., Murai, T., Okada, T., Hayashi, A., & Sengoku, A. (2001). Frontal midline theta rhythm is correlated with cardiac autonomic activities during the performance of an attention demanding meditation procedure. *Cognitive Brain Research*, *11*(2), 281–287. https://doi.org/10.1016/S0926-6410(00)00086-0

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, 621-647. https://doi.org/10.1146/annurev.psych.093008.131123

Kutas, M., & Hillyard, S. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203–205. https://doi.org/10.1126/science.7350657

Lambert, W. E., & Rawlings, C. (1969). Bilingual processing of mixed-language associative networks. *Journal of Verbal Learning and Verbal Behavior*, *8*(5), 604–609. https://doi.org/10.1016/S0022-5371(69)80111-8

Lebo, C. P., Smith, M. F. W., Mosher, E. R., Jelonek, S. J., Schwind, D. R., Decker, K. E., Krusemark, H. J., & Kurz, P. L. (1994). Restaurant noise, hearing loss, and hearing aids. *Western Journal of Medicine, 161*(1), 45-49. PMC1011367.

Lev-Ari, S., van Heugten, M., & Peperkamp, S. (2017). Relative difficulty of understanding foreign accents as a marker of proficiency. *Cognitive Science*, *41*(4), 1106–1118. https://doi.org/10.1111/cogs.12394

Li, P., Zhang, F., Yu, A., & Zhao, X. (2020). Language History Questionnaire (LHQ3): An enhanced tool for assessing multilingual experience. *Bilingualism: Language and Cognition*, *23*(5), 938–944. https://doi.org/10.1017/S1366728918001153

Loebach, J. L., Pisoni, D. B., & Svirsky, M. A. (2009). Transfer of auditory perceptual learning with spectrally reduced speech to speech and nonspeech tasks: implications for cochlear implants. *Ear and Hearing, 30*(6), 662-674. https://doi.org/10.1097/AUD.0b013e3181b9c92d

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00213

Luck, S. J., & Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't): How to get significant effects. *Psychophysiology*, *54*(1), 146–157. https://doi.org/10.1111/psyp.12639

Makeig, S. (1993). Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones. *Electroencephalography and Clinical Neurophysiology*, *86*(4), 283–293. https://doi.org/10.1016/0013-4694(93)90110-H

Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes*, *27*(7–8), 953–978. https://doi.org/10.1080/01690965.2012.705006

Maurer, U., Brem, S., Liechti, M., Maurizio, S., Michels, L., & Brandeis, D. (2015). Frontal midline theta reflects individual task performance in a working memory task. *Brain Topography*, *28*(1), 127–134. https://doi.org/10.1007/s10548-014-0361-y

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1–86. https://doi.org/10.1016/0010-0285(86)90015-0

McLaughlin, D. J., Baese-Berk, M. M., Bent, T., Borrie, S. A., & Van Engen, K. J. (2018). Coping with adversity: Individual differences in the perception of noisy and accented speech. *Attention, Perception, & Psychophysics*, *80*(6), 1559–1570. https://doi.org/10.3758/s13414-018-1537-4

Meyer, L., Obleser, J., & Friederici, A. D. (2013). Left parietal alpha enhancement during working memory-intensive sentence processing. *Cortex*, *49*(3), 711–721. https://doi.org/10.1016/j.cortex.2012.03.006

Morey, R. D., & Rouder, J. N. (2015). Bayes factor approaches for testing interval null hypotheses. *Psychological Methods, 20*(3), 307-320. https://doi.org/10.1037/met0000021

Munro, M. J., & Derwing, T. M. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech*, *38*(3), 289–306. https://doi.org/10.1177/002383099503800305

Nagle, C. L., & Baese-Berk, M. M. (2022). Advancing the state of the art in L2 speech perception-production research: Revisiting theoretical assumptions and methodological practices. *Studies in Second Language Acquisition*, *44*(2), 580–605. https://doi.org/10.1017/S0272263121000371

Nakamura-Palacios, E. M., Falçoni Júnior, A. T., Anders, Q. S., De Paula, L. D. S. P., Zottele, M. Z., Ronchete, C. F., & Lirio, P. H. C. (2023). Would frontal midline theta indicate cognitive changes induced by non-invasive brain stimulation? A mini review. *Frontiers in Human Neuroscience*, *17*, 1116890. https://doi.org/10.3389/fnhum.2023.1116890

Norris, D. (1999). The merge model: Speech perception is bottom-up. *The Journal of the Acoustical Society of America*, *106*(4), 2295–2295. https://doi.org/10.1121/1.427854

Obleser, J., & Kotz, S. A. (2011). Multiple brain signatures of integration in the comprehension of degraded speech. *NeuroImage*, *55*(2), 713–723. https://doi.org/10.1016/j.neuroimage.2010.12.020

Obleser, J., & Weisz, N. (2012). Suppressed alpha oscillations predict intelligibility of speech and its acoustic details. *Cerebral Cortex*, *22*(11), 2466–2477. https://doi.org/10.1093/cercor/bhr325

Obleser, J., Wostmann, M., Hellbernd, N., Wilsch, A., & Maess, B. (2012a). Adverse listening conditions and memory load drive a common alpha oscillatory network. *Journal of Neuroscience*, *32*(36), 12376–12383. https://doi.org/10.1523/JNEUROSCI.4908-11.2012

Obleser, J., Wostmann, M., Hellbernd, N., Wilsch, A., & Maess, B. (2012b). Adverse listening conditions and memory load drive a common alpha oscillatory network. *Journal of Neuroscience*, *32*(36), 12376–12383. https://doi.org/10.1523/JNEUROSCI.4908-11.2012

Orellana, D., Busch-Vishniac, I. J., & West, J. E. (2007). Noise in the adult emergency department of Johns Hopkins Hospital. *The Journal of the Acoustical Society of America, 121*(4), 1996-1999. https://doi.org/10.1121/1.2642309

Pallier, C., Colomé, A., & Sebastián-Gallés, N. (2001). The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science*, *12*(6), 445–449. https://doi.org/10.1111/1467-9280.00383

Peña, M., & Melloni, L. (2012). Brain oscillations during spoken sentence processing. *Journal of Cognitive Neuroscience*, *24*(5), 1149–1164. https://doi.org/10.1162/jocn_a_00144

Penolazzi, B., Angrilli, A., & Job, R. (2009). Gamma EEG activity induced by semantic violation during sentence reading. *Neuroscience Letters*, *465*(1), 74–78. https://doi.org/10.1016/j.neulet.2009.08.065

Pu, Yi, Douglas Cheyne, Yanan Sun, and Blake W. Johnson. "Theta oscillations support the interface between language and memory." NeuroImage 215 (2020): 116782. https://doi.org/10.1016/j.neuroimage.2020.116782

Rogers, C. L., Dalby, J., & Nishi, K. (2004). Effects of noise and proficiency on intelligibility of Chinese-accented English. *Language and Speech*, *47*(2), 139–154. https://doi.org/10.1177/00238309040470020201

Romero-Rivas, C., Martin, C. D., & Costa, A. (2015). Processing changes when listening to foreign-accented speech. *Frontiers in Human Neuroscience*, *9*. https://doi.org/10.3389/fnhum.2015.00167

Romero-Rivas, C., Martin, C. D., & Costa, A. (2016). Foreign-accented speech modulates linguistic anticipatory processes. *Neuropsychologia*, *85*, 245–255. https://doi.org/10.1016/j.neuropsychologia.2016.03.022

Rommers, J., Dijkstra, T., & Bastiaansen, M. (2013). Context-dependent semantic processing in the human brain: Evidence from idiom comprehension. *Journal of Cognitive Neuroscience*, *25*(5), 762–776. https://doi.org/10.1162/jocn_a_00337

Rosen, S., Souza, P., Ekelund, C., & Majeed, A. A. (2013). Listening to speech in a background of other talkers: effects of talker number and noise vocoding. *Journal of the Acoustic Society of America,* 133, 2431–2443. https://doi.org/10.1121/1.4794379

Rossi, E., & Prystauka, Y. (2020). Oscillatory brain dynamics of pronoun processing in native Spanish speakers and in late second language learners of Spanish. *Bilingualism: Language & Cognition, 23*, 964-977. https://doi.org/10.1017/S1366728919000798

Samuel, A. G., & Kraljic, T. (2009). Perceptual learning for speech. *Attention, Perception, & Psychophysics*, *71*(6), 1207–1218. https://doi.org/10.3758/APP.71.6.1207

Schiller, N. O., Boutonnet, B. P.-A., De Heer Kloots, M. L. S., Meelen, M., Ruijgrok, B., & Cheng, L. L.-S. (2020). (Not so) great expectations: Listening to foreign-accented speech reduces the brain's anticipatory processes. *Frontiers in Psychology*, *11*, 2143. https://doi.org/10.3389/fpsyg.2020.02143

Schmalz, X., Biurrun Manresa, J., & Zhang, L. (2023). What is a Bayes factor? *Psychological Methods, 28*(3), 705–718. https://doi.org/10.1037/met0000421

Schneider, J. M., Abel, A. D., Ogila, D. A., Middleton, A. E., & Maguire, M. J. (2016). Developmental differences in beta and theta power during sentence processing. *Developmental Cognitive Neuroscience, 19*, 19-30. http://dx.doi.org/10.1016/j.dcn.2016.01.001

Schneider, J. M., & Maguire, M. J. (2019). Developmental differences in the neural correlates supporting semantics and syntax during sentence processing. *Developmental Science*, *22*(4), e12782. https://doi.org/10.1111/desc.12782

Seifi Ala, T., Graversen, C., Wendt, D., Alickovic, E., Whitmer, W. M., & Lunner, T. (2020). An exploratory study of EEG alpha oscillation and pupil dilation in hearing-aid users during effortful listening to continuous speech. *PLOS ONE*, *15*(7), e0235782. https://doi.org/10.1371/journal.pone.0235782

Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, *270*(5234), 303–304. https://doi.org/10.1126/science.270.5234.303

Strauß, A., Kotz, S. A., & Obleser, J. (2013). Narrowed expectancies under degraded speech: Revisiting the N400. *Journal of Cognitive Neuroscience*, *25*(8), 1383–1395. https://doi.org/10.1162/jocn_a_00389

Sumner, M. (2011). The role of variation in the perception of accented speech. *Cognition*, *119*(1), 131–136. https://doi.org/10.1016/j.cognition.2010.10.018

To, W. M., & Chung, A. W. L. (2015). Restaurant noise: Levels and temporal characteristics. *Noise & Vibration Worldwide, 46*(8), 11-17. https://doi.org/10.1260/0957-4565.46.8.11

Töllner, T., Wang, Y., Makeig, S., Müller, H. J., Jung, T.-P., & Gramann, K. (2017). Two independent frontal midline theta oscillations during conflict detection and adaptation in a Simon-type manual reaching task. *The Journal of Neuroscience*, *37*(9), 2504–2515. https://doi.org/10.1523/JNEUROSCI.1752-16.2017

Tuladhar, A. M., Huurne, N. ter, Schoffelen, J.-M., Maris, E., Oostenveld, R., & Jensen, O. (2007). Parieto-occipital sources account for the increase in alpha activity with working memory load. *Human Brain Mapping*, *28*(8), 785–792. https://doi.org/10.1002/hbm.20306

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

Waldhauser, G. T., Blaüml, & Hanslmayer, S. (2015). Brain oscillations mediate successful suppression of unwanted memories. *Cerebral Cortex, 25(11)*, 4180-4190. https://doi.org/10.1093/cercor/bhu138

Wambacq, I., Ghanim, I., Greenfield, S., Koehnke, J., Besing, J., Chauvette, C., & Yesis, C. (2023). Neurophysiologic patterns of semantic processing of accented speech. *Journal of Neurolinguistics*, *65*, 101117. https://doi.org/10.1016/j.jneuroling.2022.101117

Wang, L., Jensen, O., Van Den Brink, D., Weder, N., Schoffelen, J.-M., Magyari, L., Hagoort, P., & Bastiaansen, M. (2012). Beta oscillations relate to the N400m during language comprehension. *Human Brain Mapping*, *33*(12), 2898–2912. https://doi.org/10.1002/hbm.21410

Wang, L., Zhu, Z., & Bastiaansen, M. (2012). Integration or predictability? A further specification of the functional role of gamma oscillations in language comprehension. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00187

Weber, A., Betta, A. M. D., & 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. https://doi.org/10.1016/j.wocn.2014.05.002

Wester, F., Gilbers, D., & Lowie, W. (2007). Substitution of dental fricatives in English by Dutch L2 speakers. *Language Sciences*, *29*(2–3), 477–491. https://doi.org/10.1016/j.langsci.2006.12.029

Wilson, E. O., & Spaulding, T. J. (2010). Effects of noise and speech intelligibility on listener Comprehension and processing time of Korean-accented English. *Journal of Speech, Language, and Hearing Research*, *53*(6), 1543–1554. https://doi.org/10.1044/1092-4388(2010/09-0100)

Whitton, J. P., Hancock, K. E., Shannon, J. M., & Polley, D. B. (2017). Audiomotor perceptual training enhances speech intelligibility in background noise. *Current Biology, 27*(21), 3237-3247. https://doi.org/10.1016/j.cub.2017.09.014

Wöstmann, M., Lim, S.-J., & Obleser, J. (2017). The human neural alpha response to speech is a proxy of attentional control. *Cerebral Cortex*, *27*(6), 3307–3317. https://doi.org/10.1093/cercor/bhx074

Zhang, L., Schlaghecken, F., Harte, J., & Roberts, K. L. (2021). The influence of the type of background noise on perceptual learning of speech in noise. *Frontiers in Neuroscience, 15*, 646137. https://doi.org/10.3389/fnins.2021.646137

Acknowledgements

This research was supported by NSF BCS-2041204 to Janet G. van Hell and Yushuang Liu and NSF BCS-2041264, NSF DGE NRT 2125865, DUE IUSE-1726811, and NSF OISE 1545900 to Janet G. van Hell. We thank Dr. Rafał Jończyk, Zepeng Zhu, Dr. Holly Zaharchuk, Mary Henderson, Erli Tang, and Panyang Li for assistance with this project. We also thank Dr. Tessa Bent for generously providing us with the multi-talker babble recordings.