Neuropsychologia 170 (2022) 108224

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Examining the context benefit in older adults: A combined behavioral-electrophysiologic word identification study

Rebecca E. Bieber [a,*], Christian Brodbeck [b], Samira Anderson [a]

[a] Department of Hearing and Speech Sciences, 0100 Lefrak Hall, University of Maryland College Park, College Park MD, 20740, USA
[b] Department of Psychological Sciences, University of Connecticut, Storrs CT, 06269, USA

ARTICLE INFO

Keywords:
Aging
Context
Speech recognition
Non-native speech

ABSTRACT

When listening to degraded speech, listeners can use high-level semantic information to support recognition. The literature contains conflicting findings regarding older listeners' ability to benefit from semantic cues in recognizing speech, relative to younger listeners. Electrophysiologic (EEG) measures of lexical access (N400) often show that semantic context does not facilitate lexical access in older listeners; in contrast, auditory behavioral studies indicate that semantic context improves speech recognition in older listeners as much as or more than in younger listeners. Many behavioral studies of aging and the context benefit have employed signal degradation or alteration, whereas this stimulus manipulation has been absent in the EEG literature, a possible reason for the inconsistencies between studies. Here we compared the context benefit as a function of age and signal type, using EEG combined with behavioral measures. Non-native accent, a common form of signal alteration which many older adults report as a challenge in daily speech recognition, was utilized for testing. The stimuli included English sentences produced by native speakers of English and Spanish, containing target words differing in cloze probability. Listeners performed a word identification task while 32-channel cortical responses were recorded. Results show that older adults' word identification performance was poorer in the low-predictability and non-native talker conditions than the younger adults', replicating earlier behavioral findings. However, older adults did not show reduction or delay in the average N400 response as compared to younger listeners, suggesting no age-related reduction in predictive processing capability. Potential sources for discrepancies in the prior literature are discussed.

1. Introduction

Semantic context supports speech recognition (Kalikow et al., 1977; Miller et al., 1951; Nittrouer and Boothroyd, 1990). For instance, words in meaningful sentences are generally recognized more accurately than words presented in isolation, and words in highly constraining sentence contexts are recognized more accurately than words in weakly constraining contexts. The literature regarding age effects on the context benefit is varied. Older adults are able to benefit from the presence of semantic context in auditory tasks requiring behavioral recognition of noisy or degraded speech (Dubno et al., 2000; Pichora-Fuller, 2008; Sheldon et al., 2008). However, a closer examination of the factors contributing to this context benefit shows that, compared to younger listeners, older adults demonstrate an exaggerated reliance on semantic context (Hartman and Hasher, 1991; Rogers and Wingfield, 2015; Sommers and Danielson, 1999), or reductions in predictive processing capacity (Federmeier et al., 2002, 2003; Wlotko et al., 2012). These variable findings could be due to the various methods used to describe the context benefit, including both behavioral and electrophysiologic outcome measures.

1.1. Electrophysiologic measures of a semantic context benefit

The N400 component is a common electrophysiologic measure used to investigate the effects of semantic context on speech recognition. The N400 component is a negative-going potential occurring around ~300–500 ms, which is thought to index the relative ease of lexical access and semantic integration (Lau et al., 2008, 2009). The magnitude of the N400 deflection reflects ease of lexical processing (Federmeier and Kutas, 1999; Lau et al., 2009): the N400 component in response to a target item which is relatively more difficult to map to a stored lexical representation will have a larger (i.e. more negative) amplitude than one

* Corresponding author.
E-mail address: rbieber@umd.edu (R.E. Bieber).

https://doi.org/10.1016/j.neuropsychologia.2022.108224
Received 28 August 2021; Received in revised form 18 February 2022; Accepted 22 March 2022
Available online 26 March 2022

in response to an item that is easier to access. Experimental conditions can be compared by examining the N400 *effect*, which is derived by calculating a difference potential: typically, a subtraction of the lexically "easier" condition from the lexically "harder" condition.

Many factors can contribute to the relative ease or difficulty of lexical access. The standard example is a comparison of conditions in which the same sentence frames are presented with different target words varying in cloze probability (a measure indicating the likelihood that the sentence frame will be completed with that word), i.e. an ***N400 effect of context***. An example of this contrast is "I like my coffee with cream and SUGAR/DOGS" (Federmeier et al., 2003; Kutas and Hillyard, 1980). A word with lower cloze probability (here, DOGS) would require relatively greater resources to access, thus resulting in an N400 component with a greater magnitude of deflection. The magnitude of the N400 component is thought to correspond with degree of semantic expectancy. The semantic or contextual information can serve to narrow a listener's expectations about upcoming lexical items. If the target item violates these expectations, the result is a need for increased processing resources, which result in the greater N400 magnitude. This has been illustrated by a number of studies demonstrating that the magnitude of the N400 component is modulated by the degree of target word predictability within a sentence or in relation to a semantic prime (Federmeier et al., 2010; Federmeier and Kutas, 1999; Kutas and Hillyard, 1984).

1.2. Lexical access and signal type

Another potential contributor to the difficulty of lexical access is clarity of the target signal. Degradations or alterations to signal quality are thought to impact the process of lexical activation, and thus influence N400 amplitude. These changes to the signal can result from naturally-occurring sources (e.g. non-native accent), or artificial manipulations (e.g. time-compression, noise-vocoding). Influences of signal alteration have also been demonstrated for event-related potential (ERP) components reflecting earlier stages of the speech recognition process, i.e. the N1 and P2 components, corresponding to auditory object detection and feature extraction (Anderson et al., 2020; Straus et al., 2013).

A small number of studies have examined the effect of non-native talker status on N400 amplitude in the absence of explicit context or predictability manipulations: i.e., an ***N400 effect of talker*** (Goslin et al., 2012; Romero-Rivas et al., 2015). In these studies, N400 components were significantly larger in response to non-native speech, suggesting that acoustic alterations imposed by a non-native accent had not been fully normalized by the listeners in early stages of processing, and thus still had an influence at the point of lexical access.

1.3. Lexical access, signal type, and semantic context

A number of studies have examined the interactions of signal alteration and semantic context on the N400 effect, i.e. the interaction of ***N400 effects of context and effects of talker***. Overall, these studies paint an unclear picture of the effect of non-native speech on a semantically elicited N400 effect. In studies of the combinatory effects of signal alteration and semantic content on the N400 effect, purely spectral changes to the signal such as low-pass filtering and noise-vocoding seem to diminish the N400 context effect (Aydelott et al., 2006; Straus et al., 2013). These findings suggest that predictive processing becomes constrained under the limitations of a distorted signal. That is, while listeners can use contextual information to generate predictions in both degraded and non-degraded listening conditions, the advantage afforded by a high-predictability target is reduced if the sentence frame has been acoustically degraded.

Non-native talker accent, which constitutes a more global, spectro-temporal change to the signal does not seem to affect the N400 response in a predictable manner. In the studies which use non-native speech, findings include reduction (Grey and van Hell, 2017), magnification (Romero-Rivas et al., 2015), and no change to the magnitude of the response (Hanulíková and Weber, 2012), with two of three studies reporting a broader distribution of the response across electrodes (Hanulíková et al., 2012; Romero-Rivas et al., 2015). This finding of a more broadly distributed response could be interpreted to reflect a recruitment of additional cognitive resources for processing non-native speech.

1.4. Auditory object formation and non-native speech

Romero-Rivas et al. (2015) also explored the effects of non-native accent on an earlier neural response, the P200. This response, also known as the P2, is a positive-going deflection that occurs around 200 ms and is understood to reflect the early stages of auditory processing, such as auditory feature detection and object formation (Reinke et al., 2003; Tremblay et al., 2001). Enhancements in the amplitude of the P200 have also been seen following auditory training or increased exposure to sound through short-term use of hearing aid amplification (Atienza et al., 2002; Karawani et al., 2018; Tremblay et al., 2001). In the Romero-Rivas et al. (2015) study, young adult listeners showed reduced P200 amplitudes in response to non-native as compared to native speech. This finding was interpreted to indicate a greater difficulty in processing the acoustic features of non-native speech, occurring even before higher-level lexical processing. This talker effect on P200 amplitude remained constant over the course of the experiment.

1.5. Lexical access, semantic context, and aging

ERP studies of visual and auditory sentence recognition suggest that older adults may be less effective at predictive processing, and less able to take advantage of a rich semantic context (Federmeier et al., 2002, 2003; Wlotko et al., 2012). Older adults show a delay in the peak latency of the N400 effect for words occurring in sentence-final positions, or following semantic primes (Federmeier et al., 2010; Federmeier and Kutas, 2005). N400 amplitudes are similar between older adults and young adults on words that have low cloze probability or are semantically incongruous, but older adults fail to show facilitation in the corresponding high cloze or high typicality conditions (Federmeier et al., 2010; Federmeier and Kutas, 2005).

Collectively, these findings indicate that older adults are less efficient and effective at making use of semantic context to generate predictions about incoming stimuli. Essentially, the literature suggests that older adults are both less efficient and effective at making use of semantic context in otherwise unchallenging listening environments, despite demonstrating a strong bias towards reliance on context for recognition. However, much behavioral literature suggests that older adults benefit equally or more from contextual information than younger adults (Dubno et al., 2000; Goy et al., 2013; Pichora-Fuller et al., 1995; Sheldon et al., 2008; Sommers and Danielson, 1999). While behavioral tasks such as repetition or transcription are valuable in illuminating the final product of perception, repetition-based measures are more limited in their ability to answer questions about the processes underlying speech recognition and perception. Objective methods such as eye-gaze measures and electrophysiology have therefore been critical in expanding the understanding of speech recognition processes from cochlea to cortex. These objective measures may be especially useful when investigating the nature of age-related declines in auditory performance. In the present study, both behavioral and EEG measures are combined in order to elucidate the level of processing at which age effects manifest in speech recognition.

1.6. Short-term changes in performance

The time-course of performance within a condition or experiment may also shed light on some of the conflicting prior findings regarding the context benefit in older versus younger adults. An example of this strategy can be seen in the study completed by Romero-Rivas et al.

(2015), which revealed that in younger adults, the N400 effect of talker decreased in magnitude over the course of the experiment. The authors interpreted this finding to reflect rapid adaptation to the non-native speech signal and increasing ease of lexical access with additional exposure to the non-native speech. Examination of this time-course data can provide information about this rapid adaptation phenomenon, which is thought to reflect the early processes of perceptual learning and support recognition under challenging circumstances (Banai and Lavie, 2020). As seen in Romero-Rivas et al.'s (2015) study, it is possible that some effects or interactions are present during only a portion of trials and shift as listeners adapt to the stimuli; looking at average data for entire conditions may mask these findings. In the present study, time-course data for both behavioral and electrophysiologic findings are analyzed.

1.7. Individual characteristics

Listener-related factors independent of age may also influence the effects of context and talker accent on speech recognition. One such factor which may influence speech recognition is the individual's cognitive capacity. The relationship between cognitive abilities and speech recognition ability has been explored extensively, though there are still significant gaps in knowledge. The Ease of Language Understanding (ELU) model (Rönnberg et al., 2008), a well-accepted model of speech recognition, posits that working memory capacity plays an important role in facilitating speech understanding in challenging environments. Working memory represents the capacity to store and manipulate information, and often emerges as a significant predictor of individual performance for speech recognition, including speech in noise (Akeroyd, 2008; Anderson et al., 2013; Füllgrabe et al., 2015) as well as non-native speech (Banks et al., 2015; Lev-Ari, 2014). The ELU model has been updated in recent years to include consideration of other aspects of executive function that are critical for speech recognition, including inhibition (Rönnberg et al., 2013). Measures of inhibition correlate with recognition of speech in the presence of competing talkers and in challenging environments (Dey and Sommers, 2015; Janse, 2012; Sommers and Danielson, 1999). In this study, individual characteristics are tested for their contribution to the various aspects of the speech recognition process.

1.8. Present study

The goal of this study was to combine behavioral and ERP methodologies to evaluate the interactions of talker accent and predictability on speech processing, and to examine any age-related changes in the context benefit. In this study, neural processing and speech recognition are compared for target words that either have high or low cloze probability based on a carrier sentence. The stimuli are produced by a native and a non-native speaker of English. In order to comprehensively examine the effects of aging, context, and talker native language on speech processing, event-related potentials were measured in response to the stimuli, and listeners reported the target word after each sentence.

The study was designed to answer the following research questions:

1. Do both older and younger adults benefit from the presence of supportive semantic context during recognition of non-native speech?

We look for evidence of 'benefit' as follows. For behavioral measures, improved speech recognition for high vs low predictability sentences and faster reaction times to targets in high vs low predictability contexts would both indicate a context benefit. In EEG measures, a smaller negative deflection in the N400 time range in response to target words contained in high predictability vs low predictability sentences would indicate a facilitatory effect of context.

2. For both behavioral and EEG measures, is the degree of benefit similar for the non-native talker's speech as compared to the native speech?

A discrepancy between the size of the context benefit for native vs non-native speech would indicate that the presence of a naturalistic acoustic alteration to the speech signal impacted the speech recognition process in some way. If the context benefit were reduced in non-native speech, this might indicate that the alterations to the bottom-up acoustic information impeded predictive processing in the high predictability conditions. If the context benefit were exaggerated in the non-native condition, this might suggest that listeners rely more heavily on top-down processes to facilitate recognition in acoustically challenging environments.

3. If an alteration in benefit from supportive context when listening to non-native speech is observed, is it similar between older and younger adults?

It was anticipated that both younger and older adults would show N400 effects of context in both the native and non-native speech conditions, given prior findings of context-elicited N400 effects under conditions of mild signal degradation (Straus et al., 2013). Should older adults demonstrate a reduced N400 effect of context, the electrophysiologic responses would allow for a determination of whether this age effect arises from an inability to benefit from rich semantic context, or from an exacerbated detriment of processing non-native speech. Trial-by-trial time-course data were expected to reveal reductions in the effects of talker over time in younger adults, consistent with the findings of Romero-Rivas et al. (2015). Evidence of rapid adaptation was expected to be delayed and/or reduced for older adults (Adank and Janse, 2010; Bieber and Gordon-Salant, 2017).

2. Method

2.1. Participants

The participants for this study included two groups of 15 listeners, including younger listeners with normal hearing (YNH) and older listeners with normal hearing (ONH); see details in Table 1. Normal hearing is defined as pure-tone thresholds of ≤25 dB HL from 250 to 4000 Hz. Listeners reporting a history of middle ear disease or neurologic impairment were excluded from participation. Prior to testing, all listeners also completed a screening test for mild cognitive impairment (MoCA; Nasreddine et al., 2005). Listeners who did not fit the hearing-related criteria or pass the MoCA with a score of ≥26 were excluded from participation. Additionally, all listeners were required to have at least a high school education, to speak only American English as their first language, and to report no languages other than English spoken in the home before the age of 7. Listeners were also queried regarding their language history and prior exposure to non-native speech.

Table 1
Participant demographics.

Listener Group	N	Age in years (mean, SD)	Right ear pure tone average (PTA) in dB HL (mean, SD)	Age-corrected LSWMT score (mean, SD)	Age-corrected Flanker score (mean, SD)
YNH	15	21.2 (2.57)	7.5 (5.98)	103.57 (10.55)	105.86 (17.13)
ONH	15	64.39 (4.94)	10.12 (3.78)	101.43 (23.73)	96.27 (12.09)

2.2. Stimuli and procedure

2.2.1. Stimuli

Stimuli for this experiment included 100 high predictability (HP) and 100 matched low predictability (LP) sentences from the Revised Speech-In-Noise (R–SPIN) corpus (Bilger et al., 1984), recorded by two male talkers. One native speaker of English (NE) and one native speaker of Spanish (NS) were recruited from the University of Maryland community. Both talkers were males between 30 and 35 years of age. The NS talker's country of origin was Peru, and he was rated to have a moderately strong accent when speaking in English (5.74/9; Atagi and Bent, 2011). The recordings were made using a Shure MS48 microphone and a Marantz Professional PMD661 Handheld Solid State Recorder. Stimuli were spliced from the raw recordings using Adobe Audition 2018, and equalized for root-mean-square (RMS) amplitude using Praat (Boersma and Weenink, 2019). A 1000 Hz calibration tone that was equal in RMS level to the sentence stimuli was generated in Praat. These HP and LP R–SPIN stimuli were selected for their design in pairing monosyllabic target words with high and low predictability contexts, allowing for examination of sentential semantic context on identical target words. The sentences are phonetically balanced and controlled for uniformity in length. Use of these stimuli also allows for comparison with the prior behavioral studies that used this corpus (Dubno et al., 2000; Pichora-Fuller, 2008; Pichora-Fuller et al., 1995a; Sheldon et al., 2008).

2.2.2. Procedure

A total of four conditions was evaluated: High predictability, native English talker (HP, NE); low predictability, native English talker (LP, NE); high predictability, native Spanish talker (HP, NS); low predictability, native Spanish talker (LP, NS). Each listener heard 50 sentences per condition. NE and NS items were presented in separate blocks, but the HP and LP items were presented randomly within each block, resulting in one block of 100 trials per talker. Each listener heard each target word only once, and the assignment of target words to HP/LP and NE/NS was randomized across participants.

Stimulus presentation and behavioral response collection were completed using Presentation software (Neurobehavioral Systems, Berkeley, CA). Stimuli were presented monaurally to the right ear at 75 dB SPL via an ER-1 insert earphone (Etymotic Research, Elk Grove Village, IL). Each trial included the following: a fixation screen to prompt the participants to listen (500 ms), auditory presentation of a sentence, and a 3-s long response window. The response screen visually presented a closed set of 6 options, including the target word and 5 foils, which were minimal pairs to the target words. The participants' task was to select the target item as quickly and accurately as possible by pressing one of six buttons. Participants used a keyboard to select which word was heard, allowing for collection of both response accuracy and reaction time (RT). The key-press response also initiated the subsequent trial. Breaks were built into each list to allow time for eyeblinks and to ensure comfort. Prior to initiating the experiment, each listener completed a practice list of eight sentences from a different corpus produced by a NE talker who was not used otherwise during the experiment. The purpose of this practice list was to familiarize the listener with the task and use of the response keyboard; listeners were given the option to repeat the familiarization list if they needed additional practice before beginning the experiment.

Following the listening experiment, all participants completed tasks from the NIH Cognitive Toolbox, including the Flanker Task and the List Sorting Working Memory Task (Weintraub et al., 2013). In the Flanker Task, which measures inhibitory control, participants are asked to respond to a target image which is flanked by two congruent or incongruent images. Flanker scores are calculated by comparing the performance on the two types of trials (congruent and incongruent). The List Sorting Working Memory Task (LSWMT) requires participants to both recall and sort a list of items which is presented both visually and auditorily. LSWMT scores are calculated based on the number of correct trials. This task requires listeners to retain and manipulate lexical items in order to respond appropriately; listeners who show strengths in this ability may show a stronger N400 effect in the current experimental paradigm, as the benefit of high predictability sentence frames may only be realized if the participant has retained the information that occurred earlier in the sentence and can successfully integrate the subsequent items. For both the Flanker and LSWMT measures, age-corrected scores were used in the analysis; group means can be found in Table I. For two participants, Flanker and LSWMT data were lost in the data retrieval process; these were treated as missing values in the analysis.

2.3. EEG recording and signal processing

EEG responses were recorded simultaneously to the behavioral task, at a 2048-Hz sampling frequency with an antialiasing filter (low-pass filter with a −3 dB point at one-fifth of the sampling frequency) with the Biosemi Active Two system (Biosemi B.V., Netherlands). Data were recorded with an all-pass filter using a 32-channel cap. Electrodes on the right and left earlobes (A1 and A2) served as reference electrodes, with additional electrodes placed above and beside the left eye to record eye movements. Event triggers were marked at the onset of the first word and the target word of each sentence. Data were analyzed offline with MNE-Python (Gramfort et al., 2014) and Eelbrain (Brodbeck et al., 2021). Responses were band-pass filtered offline using a zero-phase FIR filter with half-amplitude cutoff frequencies of 0.1 and 40 Hz for analysis. Correction of artifacts such as eyeblinks and heartbeats was completed using independent component analysis (ICA) with all participants. Following ICA, responses were separated into epochs from -200 - 1000 ms relative to the timepoints of interest: start of the first word and start of the target word. All epochs were baseline corrected for the 200 ms period prior to the word onset. Epochs exceeding a threshold of 100 μv after ICA were automatically removed from analysis. The remaining epochs were checked visually, and channels containing excessive noise were reconstructed using spherical spline interpolation. The average number of clean epochs per participant was 177/200 for the first word, and 174/200 for the target word.

3. Analysis and results

3.1. Behavioral results

3.1.1. Analysis

Word identification accuracy scores and reaction times were each calculated per trial. Accuracy was coded as a binary outcome reflecting accurate identification of the target word (1 = correct; 0 = incorrect) and response times were measured in milliseconds as the time between the onset of the response window and the keypress response. Relative reaction times were derived for each participant, with their individual mean RT in the NE-HP condition serving as baseline; relative RT was the subtraction of individual baseline from each trial's RT.

Mixed-effects regression models were constructed for both the accuracy and RT measures. The models were constructed using the model-building procedure recommended by Hox et al. (2010). In this procedure, terms are included sequentially in the model with likelihood ratio testing used to determine whether each term significantly improved model fit. For both the accuracy and the RT models, the following terms and their interactions were evaluated: listener group (2 levels, dummy coded, reference level = YNH); predictability (2 levels, dummy coded, reference level = HP), talker accent (2 levels, dummy coded, reference level = Native English); LSWMT score (continuous predictor); Flanker score (continuous predictor). Additionally, based on visualizations of the raw data, first and second-order orthogonal polynomial time terms were created in order to independently describe the possible linear and non-linear features of the performance curve (Mirman, 2014), and were evaluated for their contribution to model fit. Random intercepts for participant and token, as well random slopes for talker accent and predictability, were evaluated for significant contribution to the models. For models with random slopes which resulted in singular fits, the random slopes were excluded from the final model to avoid overfitting.

3.1.2. Results: Word identification accuracy

The final generalized linear mixed effects model selected todescribe the word identification accuracy data was $Accuracy \sim Predictability^{*} Group^{*} Talker + TrialNumber_{linear} + (Talker|Subject) + (Talker|Word)$. The full model output is included in Table 2. See Fig. 1 for visualizations of the interactions of talker, predictability, and listener age.

The significant three-way interaction between predictability, group, and talker resulted from a larger predictability effect for the native Spanish talker than for the native English talker, which was amplified for the ONH listeners as compared to the YNH listeners ($\beta = 0.91$, SE $= 0.45$, $z = 2.04$, $p < .05$). Further examination of the interaction reveals that the ONH listeners showed predictability effects for both the native English ($\beta = 0.71$, SE $= 0.22$, $z = 3.2$, $p < .01$) and native Spanish ($\beta = 1.62$, SE $= 0.22$, $z = 7.43$, $p < .001$) talkers, whereas the YNH listeners did not show a predictability effect for the native English talker ($\beta = -0.19$, SE $= 0.25$, $z = -0.74$, $p = .46$). Within the high predictability condition, neither the YNH ($\beta = -0.46$, SE $= 0.31$, $z = -1.51$, $p = .13$) nor ONH ($\beta = 0.01$, SE $= 0.3$, $z = 0.04$, $p = .97$) listeners showed an effect of talker, suggesting that the presence of non-native talker accent alone did not significantly reduce performance, when supportive semantic context was available.

The significant main effect of Trial number indicated that performance increased significantly across trials ($\beta = 1.41$, SE $= 0.54$, $z = 2.6$, $p < .01$). The quadratic time term was not found to contribute significantly to model fit ($p > .05$), and the linear time term was not found to interact significantly with any of the other fixed predictors ($p > .05$, all comparisons), suggesting that the performance increases were similar across all conditions. Listeners' performance on the cognitive tasks was also tested for contribution to model fit, but none were found to significantly improve the model ($p > .05$, all comparisons).

In summary, older adult listeners had lower word identification accuracy for target words in low-predictability contexts, especially for stimuli spoken by a Spanish-accented talker. Younger adults also showed this predictability effect for the Spanish-accented stimuli, but not for unaccented stimuli. All listeners improved their word identification accuracy over time, regardless of talker or predictability.

Table 2
Word identification accuracy by trial.

Predictors	Word Identification Accuracy			
	Odds Ratios	std. Error	z-value	P
(Intercept)	26.53	7.45	11.68	<.001
Predictability [LP]	1.20	0.30	0.74	.46
Talker [Native Spanish]	1.59	0.49	1.50	.13
$TrialNumber_{linear}$	4.11	2.24	2.60	<.01
Group [ONH]	1.08	0.41	0.21	.83
Predictability [LP] * Talker [Native Spanish]	0.16	0.05	−5.43	<.001
Predictability [LP] * Group [ONH]	0.41	0.13	−2.71	<.01
Talker [Native Spanish] * Group [ONH]	0.62	0.24	−1.23	.22
(Predictability [LP] * Talker [Native Spanish]) * Group [ONH]	2.48	1.11	2.04	<.05
Random Effects				
σ^2	3.29			
$\tau_{00\ Word}$	0.29			
$\tau_{00\ Subject}$	0.58			
$\tau_{11\ Word.TalkerNative\ Spanish}$	1.18			
$\tau_{11\ Subject.TalkerNative\ Spanish}$	0.10			
$\rho_{01\ Word}$	−0.15			
$\rho_{01\ Subject}$	−0.93			
ICC	0.27			
$N_{Subject}$	30			
N_{Word}	200			
Observations	5999			
Marginal R^2/Conditional R^2	0.095/0.336			

3.1.3. Results: Reaction times

Relative reaction times of correct trials were fitted to a linear mixed effects regression following the procedures described in section 3.1.1. The final model selected to describe the relative reaction time data was: $RT_{rel} \sim Predictability^{*} Talker + (TrialNumber_{linear} {}^{*}Group) + (1|Subject) + (1|Word)$. The relative reaction times are visualized in Fig. 2, and the full model summary is found in Table 3.

The significant interaction of Predictability and Talker reflects the presence of a context effect for the NS talker, but not for the NE talker ($\beta = 1064.26$, SE $= 258.64$, $t = 4.12$, $p < .001$). Relative reaction times are slower for the NS talker in the low predictability sentences, consistent with increased effort associated with the word identification process in this condition. The interaction of trial number and listener group shows that the relative RT increased over time for the YNH listeners, but decreased over time for the ONH listeners, for both talkers: the effect of trial number on RT is significantly more negative for the ONH listeners than for the YNH listeners ($\beta = -31.37$, SE $= 8.79$, $z = -3.57$, $p < .001$). Listeners' performance on the cognitive tasks was also tested for contribution to model fit, but none were found to significantly improve the model ($p > .05$, all comparisons).

Thus, relative reaction times were slower for low predictability than high predictability stimuli, but only for the stimuli produced by the Native Spanish talker. Younger listeners showed steadily increasing relative RTs over time, while ONH listeners' RTs decreased with additional listening in each condition.

3.2. Average EEG responses

The ERPs of interest within the response to the first word of the sentence were the P200 and the N400. The P200 response to the first word of the sentence provides information about the effect of aging and talker language background on processing of acoustic information, while the N400 response to the first word of the sentence can be used to examine the relative ease of lexical processing for the two talkers, absent any context manipulations.

3.2.1. Analysis: First word

Both latency and amplitude of the P200 were examined. The average EEG response from the Cz and Fz sensors for each individual subject was plotted and P200 peaks were marked by hand by the first and third authors. This strategy was used rather than analyzing a pre-determined time window due to the nature of the stimuli; as the sentence onsets were not phonemically uniform, the averaged responses were broader and less distinct than the typical P200 elicited by uniform tones or speech syllables. In addition, examination of the grand averaged waveforms confirmed our hypothesis that there would be age-related latency differences in the P200, which would necessitate a very broad analysis window. The individual P200 latencies were used to calculate individual P200 amplitudes for each trial; a window of 50 ms around each individual's peak was used to generate the average P200 amplitude across Cz and Fz for each trial.

The individualized average P200 latencies were analyzed using a two-way ANOVA including age group and talker as predictor variables. P200 amplitudes were then analyzed using linear mixed-effects regression model, to account for the multiple measures per participant. A model was constructed using the model-building procedures recommended by Hox et al. (2010) as described in section 3.1.1, with evaluation of the following terms and their interactions for contribution to explaining variance in P200 amplitudes: listener group (2 levels, dummy coded, reference level = YNH); predictability (2 levels, dummy coded, reference level = HP), talker accent (2 levels, dummy coded, reference level = Native English); LSWMT score (continuous predictor); Flanker score (continuous predictor); trial number (continuous predictor).

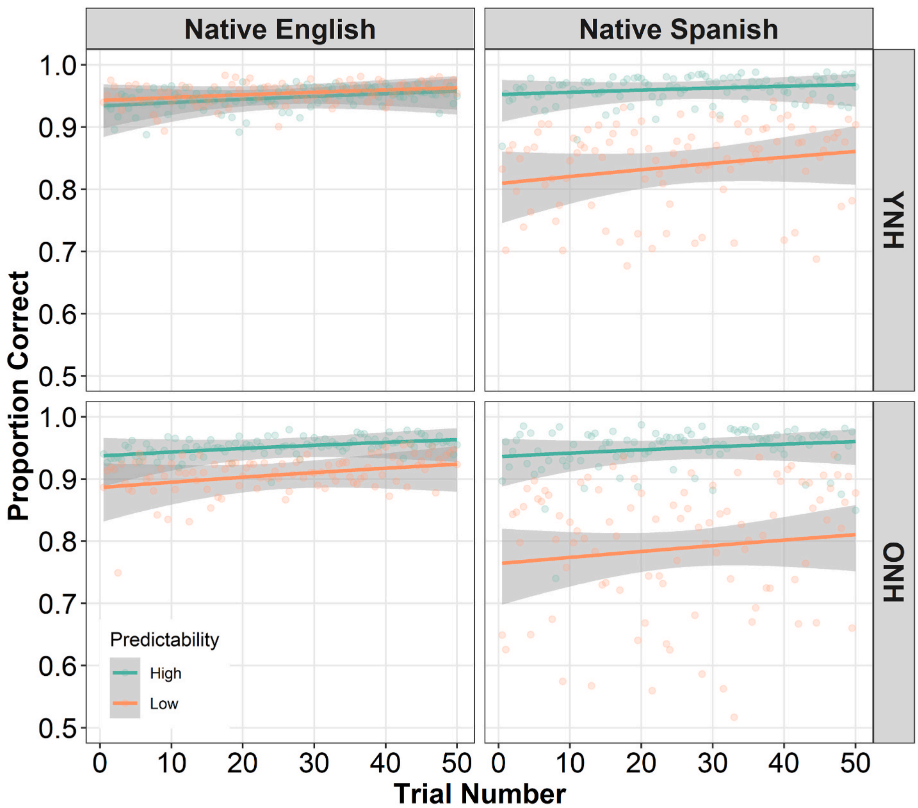

Fig. 1. Word identification accuracy is more impacted by context, together with talker accent, in older adults than in younger adults. Performance also improves as a function of trial number, regardless of listener group, talker type, or sentence predictability. Word identification performance is plotted as a function of trial number. Individual points represent group means per trial; lines reflect model predicted values with shading reflecting standard error. YNH = younger normal hearing, ONH = older normal hearing.

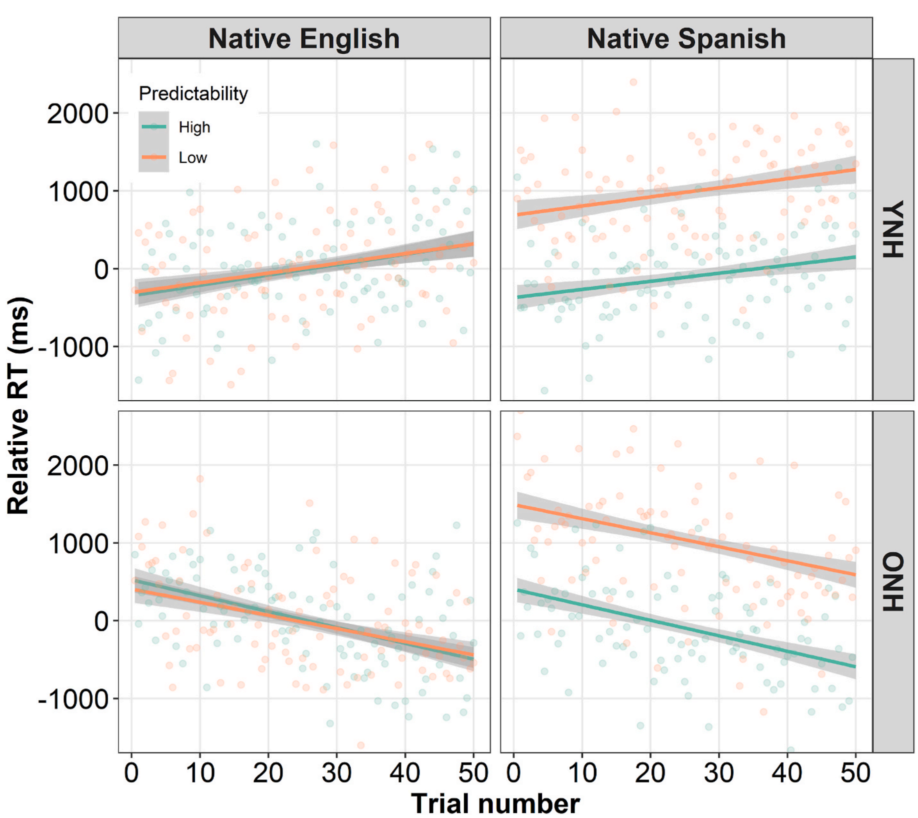

Fig. 2. Relative reaction time (RT) is slower for low predictability stimuli, but only for Spanish-accented tokens. Younger and older listeners show opposite patterns of change over trials. Relative RT is plotted as a function of trial number, separated by listener group, talker language background, and target predictability. Relative RT was calculated by subtracting the individual's mean RT in the Native English high predictability condition from their RT in each trial. Individual points represent group means per trial; lines reflect model predicted values with shading reflecting standard error. YNH = younger normal hearing, ONH = older normal hearing.

Table 3
Relative reaction time by trial.

Predictors	Relative Reaction Time			
	Estimate	std. Error	t-value	P
(Intercept)	−298.96	293.76	−1.02	.31
Predictability [LP]	69.67	184.84	0.38	.71
Talker [Native Spanish]	−97.21	176.25	−0.55	.58
Group [ONH]	818.59	363.80	2.25	<.05
$\text{TrialNumber}_{\text{linear}}$	11.03	6.39	1.73	.08
Predictability [LP] * Talker [Native Spanish]	1064.26	258.63	4.11	<.001
Group [ONH] * $\text{TrialNumber}_{\text{linear}}$	−31.37	8.78	−3.57	<.001
Random Effects				
σ^2	21811983.49			
$\tau_{00\ \text{Word}}$	1348082.24			
$\tau_{00\ \text{Subject}}$	499372.41			
ICC	0.08			
N_{Subject}	30			
N_{Word}	200			
Observations	5458			
Marginal R^2/Conditional R^2	0.010/0.088			

During visual inspection of the data, it was noted that the grand mean waveforms of the responses to sentence onset suggested that the P200 was followed by an extended negative drift, without an identifiable peak resembling the N400. In order to explore this observation, mass univariate statistics were subsequently employed to examine responses for differences in longer latency components, such as an N400-like effect of talker and any potential interactions of talker and age group (Maris and Oostenveld, 2007). This cluster-based nonparametric approach is recommended to control Type I error rates in electrophysiology experiments where precise latencies are unknown a priori, while maintaining a conservative approach to correct for multiple comparisons (Luck and Gaspelin, 2017). For each comparison, clusters were defined as contiguous regions in space and time which exceeded an *F* or *t*-value equivalent to an uncorrected *p*-value of .05, within a time window of 1.2 s from the onset of the word. To evaluate significance, the cluster-mass of these clusters was compared to a null-distribution based on 10,000 random permutations of the condition labels.

3.2.2. Results: First word

The response to the first word of the sentence is displayed in Fig. 3, with separate curves depicted for each age group and talker.

The ANOVA analysis of P200 latency confirmed a main effect of group ($F(1, 56) = 13.63, p < .001$), with the older listeners showing later P200 latency than younger adults. P200 latencies are displayed in the left-hand panel of Fig. 4. There is no main effect of talker or interaction between talker and age group on P200 latency. For P200 amplitude (Fig. 4: Right panel), the final linear mixed-effects model determined to best fit the data was: $P200_{amp} \sim Group^{*}\ Talker + (1|Subject)$. This model did include an interaction of age group and talker (β = 1.37e-06, SE = 6.14e-07, t = 2.24, $p < .05$), with younger adults showing no significant differences in amplitude between the response to the NE talker and the NS talker (β = 1.63e-07, SE = 4.40e-07, *z-ratio* = 0.37, p = .98). Older adults, however, showed a decrease in P200 amplitude in the response to the NS talker as compared to the NE talker (β = 1.54e-07, SE = 4.28e-07, *z-ratio* = 3.59, $p < .01$). Trial number was found not to have a significant impact on P200 amplitude, nor to interact with any of the fixed predictors of talker or listener group ($p > .05$, all comparisons).

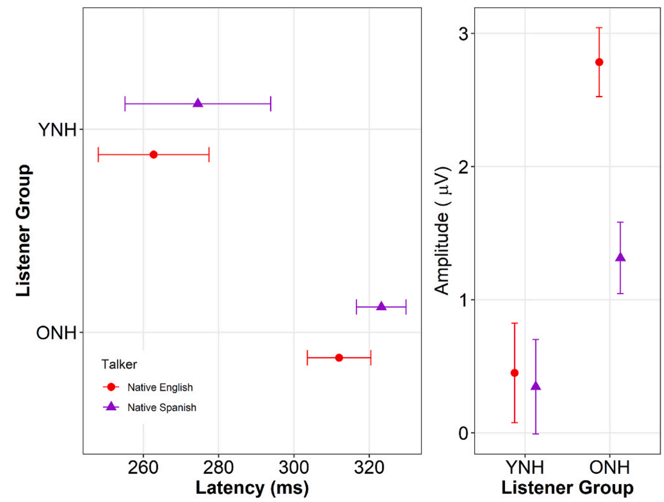

Fig. 4. Left panel: Older adults showed delayed P200 component latency relative to younger adults. Latencies are plotted separately by listener group and talker type. Right panel: Older adults show a reduction in P200 amplitude for Spanish-accented speech, despite an overall higher level of activation compared to younger adults. Average amplitudes of the P200 component across Cz and Fz electrodes are plotted separately by listener group and talker type. Error bars reflect standard error of the mean. YNH = younger normal hearing, ONH = older normal hearing.

The subsequent cluster-based analysis of later components in the first word revealed a main effect of talker beginning at 459 ms and lasting for the duration of the analysis window (459–1200 ms, $p < .01$), and a main effect of group beginning at 278 ms and lasting for the duration of the analysis window (278–1200 ms, $p < .01$). The main effect of group found by this analysis is consistent with that seen in the P200 analysis described above. The group effect is visualized in Fig. 5, and shows a greater amplitude in the ONH listener responses as compared to the YNH listeners.

The talker effect is shown in Fig. 6. The distribution of this effect (see the bottom right topographical map in Fig. 6) is consistent with an N400-like effect, which is typically distributed over the central regions. This indicates decreased ease of lexical access for the NS talker as compared to the NE talker; the effect did not interact with listener age. Similarly, trial number was found not to predict the amplitude of the N400-like component response to the first word, nor to interact with group or talker. Thus, the group and talker-related effects on these responses to the first word of the sentence are understood to remain constant across the experiment. These findings indicate that the P200 and N400 component amplitudes seen in the response to the first word of the sentence did not change significantly over the course of any listening condition.

In summary, listener age had an effect on both the latency and amplitude of the P200 component, with older adults showing delayed latencies and exaggerated amplitudes. In addition, older adults' responses were influenced by talker, with P200 amplitude decreased for the native Spanish talker. This talker effect was not seen for the younger adults.

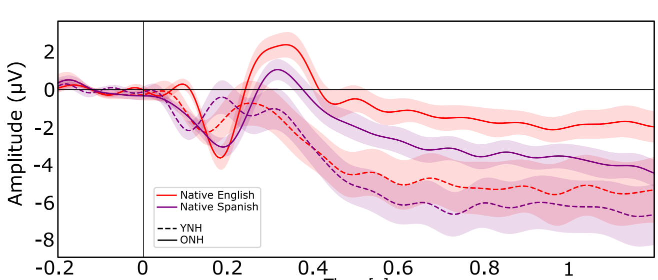

Fig. 3. Averaged ERP waveforms in response to the first word of the sentence, separated by listener group and talker language. 0 ms corresponds to the onset of the first word. The shaded region denotes standard error. Responses are averaged over the Cz and Fz sensors and filtered from 0.1 to 20 Hz for visualization only. YNH = younger normal hearing, ONH = older normal hearing.

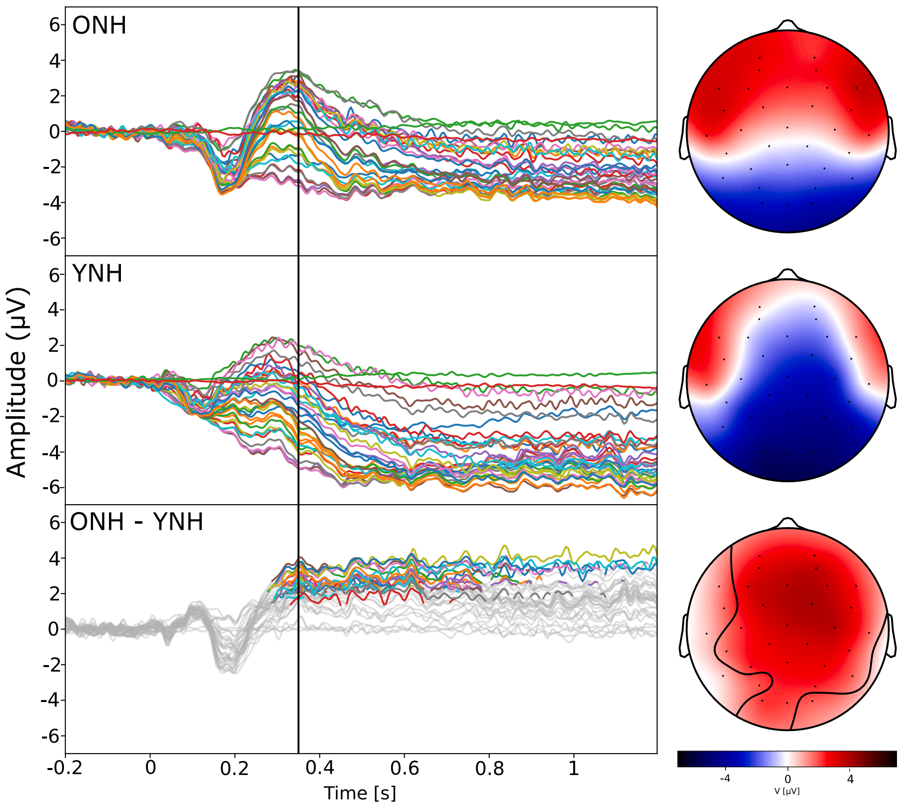

Fig. 5. Group main effect in response to the first word. Top = older normal-hearing listeners (ONH), middle = younger normal-hearing listeners (YNH), bottom = group effect (ONH – YNH). Responses are averaged across participant, talker and predictability; each line reflects the averaged response from one individual electrode channel. The solid vertical line represents time at 350 ms; the topographies at 350 ms are shown on the right. Channels that are part of the cluster with a significant difference (highlighted in the bottom panel) are bounded in the bottom topography map, indicating a centro-frontal distribution.

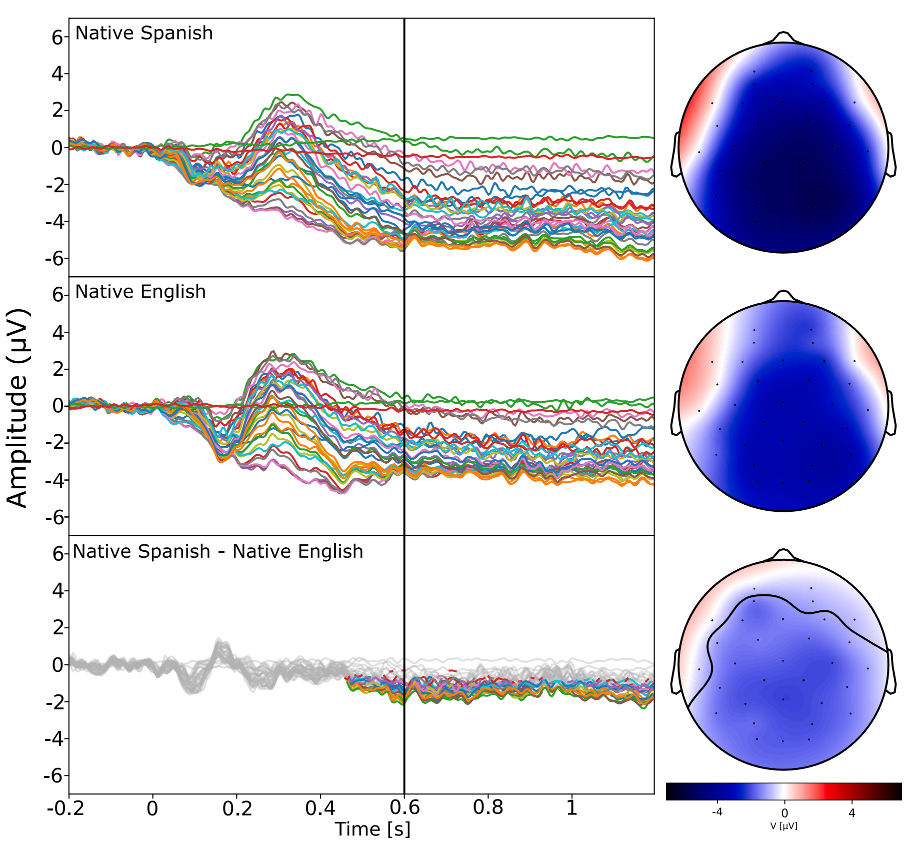

Fig. 6. Talker main effect in response to the first word. Top = Native Spanish, middle = Native English, bottom = talker effect (Native Spanish – Native English). Responses are averaged across participant, listener group and predictability; each line reflects the averaged response from one individual electrode channel. The solid vertical line represents time at 600 ms; the topographies for the talker effect at 600 ms are shown on the right. Channels that are part of the cluster with a significant difference are bounded in the bottom topography map, indicating a centro-parietal distribution.

3.2.3. Analysis: Target word

The response to the target word of the sentence was analyzed with a focus on N400 component amplitude. A time window of 300–500 ms was selected for this analysis, based on the prior literature around this component's characteristics (Federmeier et al., 2003; Federmeier and Kutas, 2005; Lau et al., 2009), as well as an examination of the grand mean waveforms. The mean amplitude across all channels between 300 and 500 ms was then analyzed using a cluster-based nonparametric approach following the procedures described in section 3.2.1. Age group, talker, and context were included as predictor variables. This analysis included only the trials in which subjects had provided a correct behavioral response (Lau et al., 2009), ensuring that the effects reflected differences in processing, rather than differences in comprehension. For the younger listeners, an average of 149 trials (77 NE) were included per subject, and for the older listeners, an average of 159 trials (80 NE) were included per subject.

A subsequent model was constructed in order to examine the N400 response to the target word of the sentence over the course of the experiment. For this model, amplitude was averaged across the Cz and Pz sensors between 300 and 500 ms for each trial. A linear mixed effects regression was fitted to these amplitudes following the model-building procedures described in section 3.1.1.

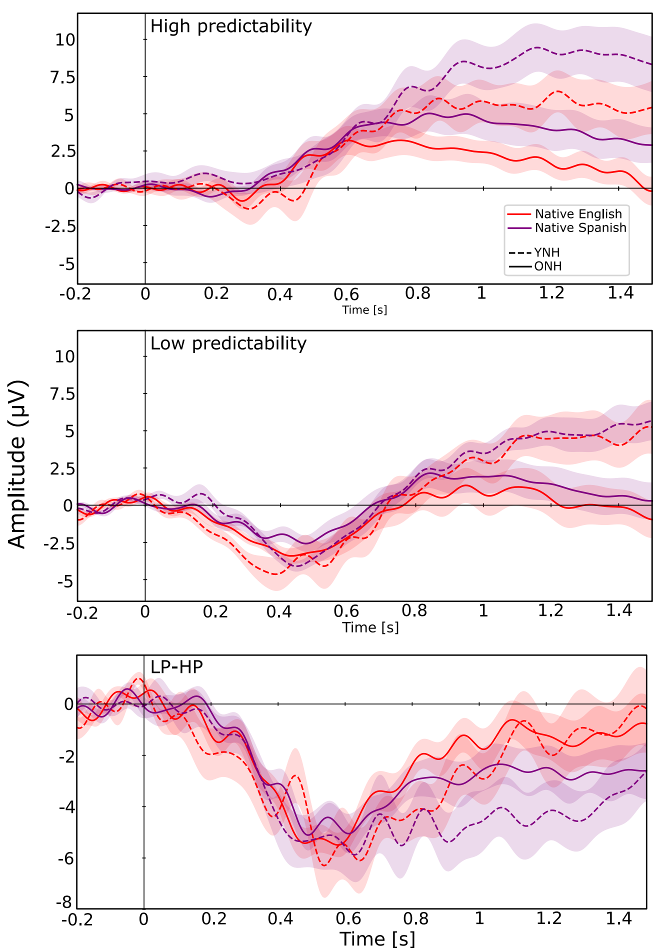

Fig. 7. Averaged ERP waveforms in response to the target word of the sentence, separated by listener group, talker language, and target predictability. Responses to target words in high-predictability (HP) and low-predictability (LP) sentences are shown in the top and middle panels, respectively. The bottom panel reflects the predictability difference wave (LP-HP), showing a robust N400 effect in each group and condition. 0 ms corresponds to the onset of the target word. The shaded region denotes standard error. Only trials with correct behavioral responses are included. Responses are averaged over the Cz and Pz sensors and filtered from 0.1 to 20 Hz for visualization only. YNH = younger normal hearing, ONH = older normal hearing.

3.2.4. Results: Target word

In Fig. 7, average responses to the target word are displayed at sensors Cz and Pz, which typically carry strong N400 signals, for both talker/listener group combinations, as well as difference waves comparing the LP and HP conditions. The mass-univariate analysis confirmed a significant effect of predictability ($p < .001$), shown in full detail in Fig. 8.

As expected, there was a larger magnitude of deflection (i.e. more negative amplitude) for the LP sentences. The central distribution of the response is consistent with the classic N400 response. There was no main effect of group, nor did group interact with any other predictor variable ($p > .05$, all comparisons). A significant effect of talker ($p < .05$) was also observed, indicating that amplitudes were more negative for the NE talker as compared to the NS talker. This talker effect is visualized in Fig. 9, and shows a centro-frontal distribution. The effects of talker and predictability did not interact in this analysis ($p > .05$).

3.2.5. N400 timecourse

The final model selected to examine the changes in N400 amplitude over time was: $MeanN400 \sim Predictability + Talker + (TrialNumber_{linear} * Group) + (Talker * WorkingMemory) + (Predictability|Subject) + (Predictability|Word)$. The full model summary is displayed in Table 4. The main effects of predictability and talker are consistent with the findings of the multivariate analysis described above. The significant interaction of trial number and group ($\beta = -6.4e\text{-}08$, SE = 2.07e-08, $t = -3.1$, $p < .01$) indicates that the two listener groups showed different patterns of N400 component amplitude change over time. The mean N400 component amplitude decreased in magnitude (i.e. increased in absolute amplitude) over the course of trials for the YNH listeners, while the ONH listeners showed the opposite pattern. This effect is visualized in Fig. 10.

The significant interaction between talker and working memory scores is visualized in Fig. 11, and suggests that there was a stronger relationship between working memory scores and N400 component amplitude for target words produced by the NE talker than the NS talker ($\beta = 8.72e\text{-}07$, SE = 3.04e-07, $t = 2.87$, $p < .01$).

Specifically, individuals with better working memory (WM) performance had N400 components with larger magnitudes, or more negative absolute amplitudes. In order to visualize the distribution of this effect, listeners were divided into high WM and low WM groups based on a median split. The topographies of the response between 300 and 500 ms is shown for both WM groups, with separate plots for each talker, in Fig. 12.

Working memory did not interact with predictability; working memory score predicted the magnitude of the N400 component, but not the N400 effect.

In sum, YNH listeners showed decreasing N400 components for both talkers and predictability conditions with additional listening experience, while older adults did not. Working memory capacity predicted the overall magnitude of the N400 component for the NE talker more strongly than for the NS talker, regardless of listener group or sentence predictability.

4. Discussion

In this study, behavioral and electrophysiologic measures were obtained simultaneously in order to examine the effect of aging on the benefit of contextual information for recognition of native and accented speech. Responses were examined over the course of trials to determine if patterns of rapid adaptation differed across age groups. The study was motivated by discrepancies in prior literature that occurred across methodologies; various findings suggest that older listeners can benefit less (Federmeier et al., 2002, 2003; Federmeier and Kutas, 2005;

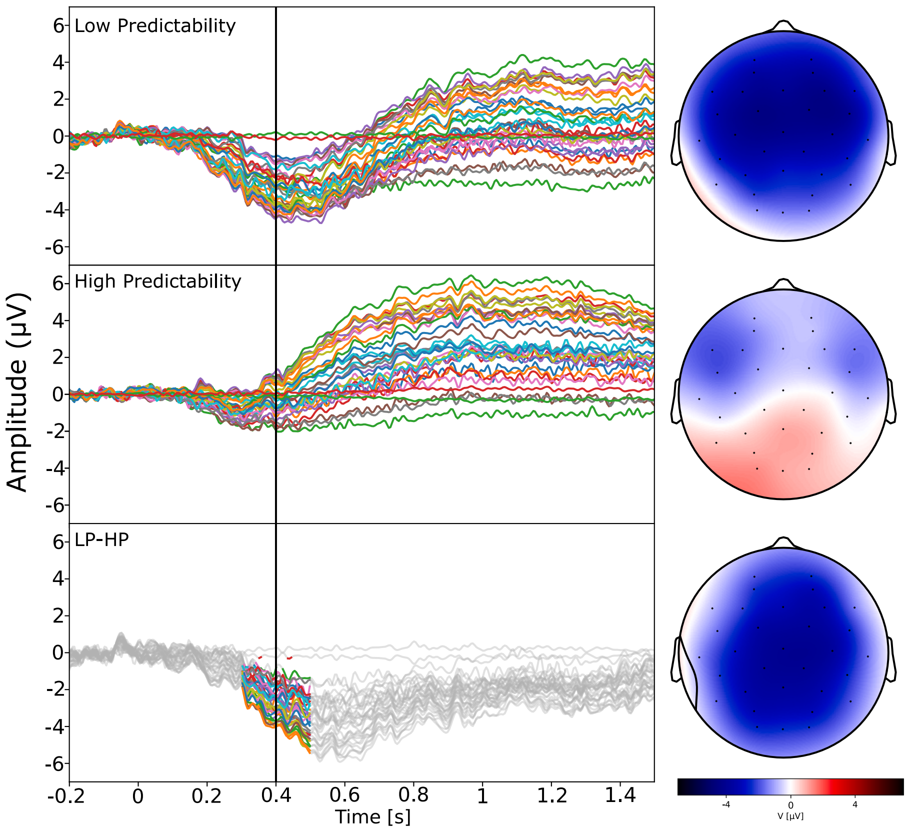

Fig. 8. Visualization of the predictability effect in response to the target word. Top = Low predictability (LP), middle = High predictability (HP), bottom = LP – HP. Responses are averaged across listener group and talker; only trials with correct behavioral responses are included. The solid vertical line represents time at 400 ms; the topographies for the predictability effect at 400 ms are shown on the right. Channels with significant difference are bounded in the bottom topography map; nearly all channels are significant in this analysis. The highlighted waveform region of 300–500 ms in the bottom panel reflects the time boundaries of the analysis window.

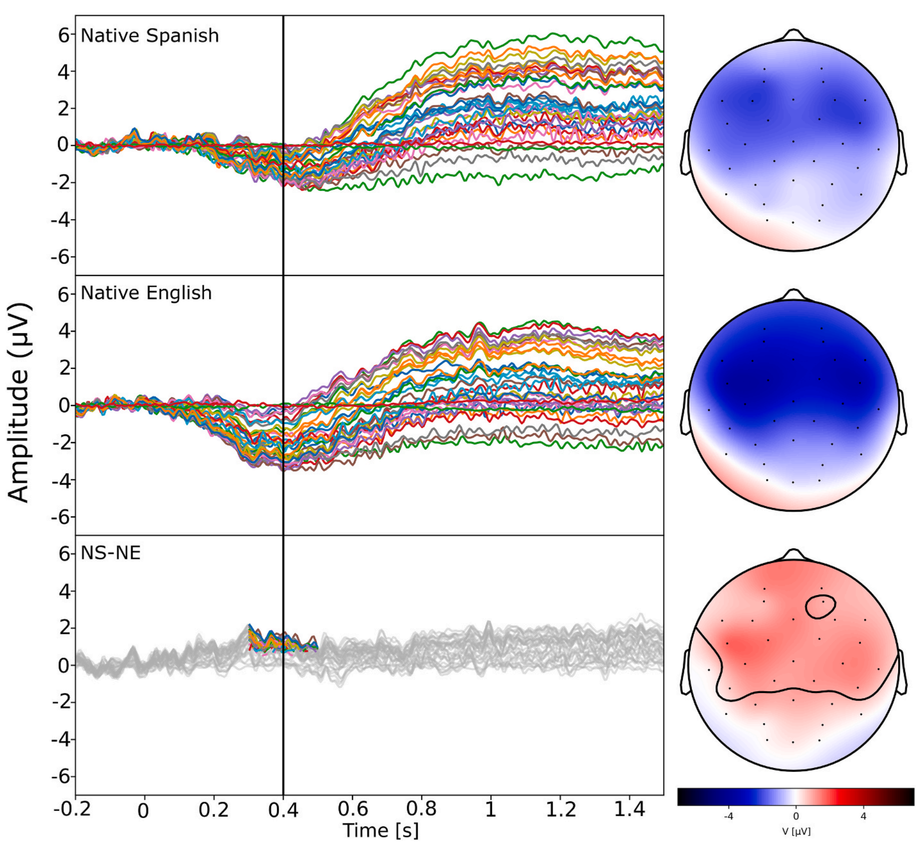

Fig. 9. Visualization of the talker effect in response to the target word. Top = Native Spanish (NS), middle = Native English (NE), bottom = NS – NE. Responses are averaged across listener group and talker. The solid vertical line represents time at 400 ms; the topographies for the predictability effect at 400 ms are shown on the right. Channels with significant difference are bounded in the bottom topography map, indicating a centro-frontal distribution of the effect. The highlighted waveform region of 300–500 ms in the bottom panel reflects the time boundaries of the analysis window.

Table 4
Target word N400 by trial.

Predictors	Mean N400 amplitude			
	Estimate	std. Error	t-value	p
(Intercept)	−1.18e-06	7.51e-07	−1.57	.12
Predictability [LP]	−3.55e-06	4.7e-07	−7.57	<.001
Talker [Native Spanish]	8.47e-07	3.01e-07	2.81	<.01
Group [ONH]	2.55e-06	8.16e-07	3.13	<.01
$TrialNumber_{linear}$	3.24e-08	1.49e-08	2.18	<.05
ListSort_Age_z	−1.04e-06	3.45e-07	−3.03	<.01
Group [ONH] * $TrialNumber_{linear}$	−6.4e-08	2.07e-08	−3.10	<.01
Talker [Native Spanish] * ListSort_Age_z	8.72e-07	3.04e-07	2.87	<.01
Random Effects				
σ^2	9.73e-11			
$\tau_{00\ word}$	3.53e-12			
$\tau_{00\ subject}$	6.77e-12			
$\tau_{11\ word.predLP}$	6.32e-12			
$\tau_{11\ subject.predLP}$	2.7e-12			
$\rho_{01\ word}$	−0.94			
$\rho_{01\ subject}$	−0.95			
ICC	0.06			
$N_{subject}$	28			
N_{word}	200			
Observations	4375			
Marginal R^2/Conditional R^2	0.042/0.102			

Schurman et al., 2014; Wlotko et al., 2012), equally (Dubno et al., 2000; Sheldon et al., 2008; Wingfield et al., 1994), or more (Goy et al., 2013; Pichora-Fuller et al., 1995b; Sommers and Danielson, 1999) than younger listeners from the presence of semantic context during otherwise challenging listening situations. Unexpectedly, the results of this study documented a dissociation between behavioral and electrophysiologic results, despite simultaneous data collection; these results suggest that the mechanisms probed by each measure are affected independently by changes to the talker- and listener-related variables examined in this study.

4.1. Benefit of context for recognition of non-native speech

In this study, 'benefit' was defined for both behavioral and EEG

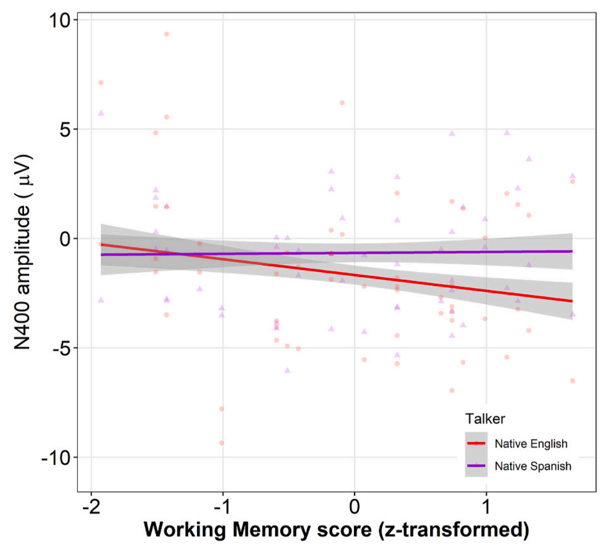

Fig. 11. Scores on the List Sorting Working Memory Task (LSWMT) are predictive of mean N400 component amplitude in response to native English speech, but not Spanish-accented stimuli. LSWMT scores were age-corrected and z-transformed for analysis. Separate regression lines are plotted for each talker, data are collapsed by group and target predictability. Points represent individual data; regression lines reflect model predicted values with shading reflecting standard error.

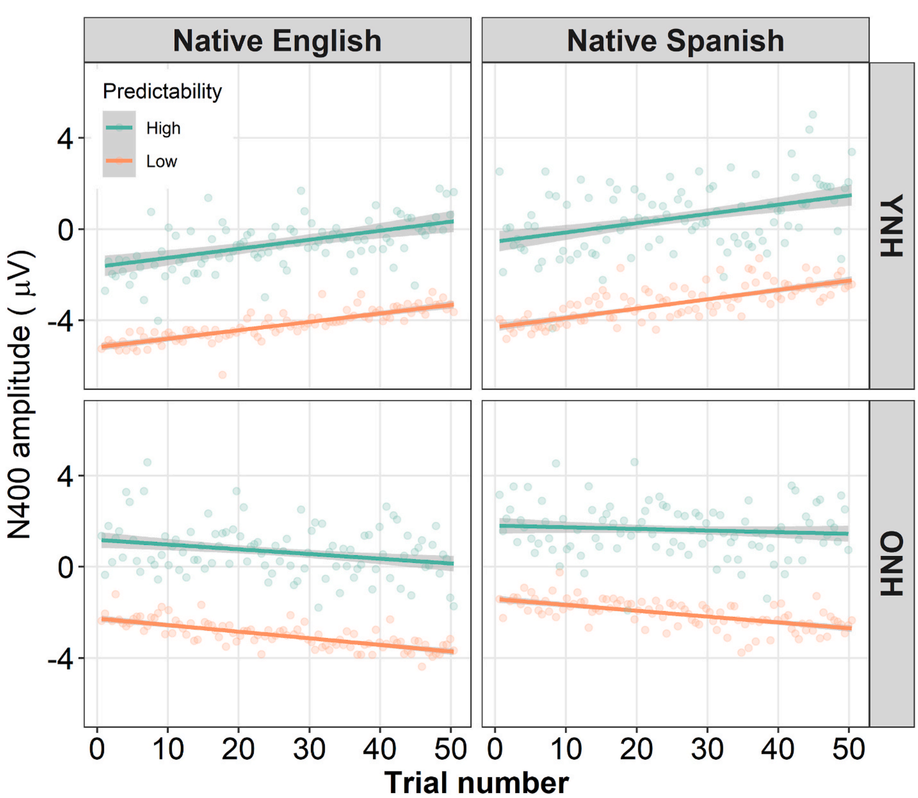

Fig. 10. N400 component amplitudes change as a function of trial, but the magnitude of the N400 effect remains constant over trials regardless of listener group or talker. Fitted values for mean N400 component amplitude in response to the target word are plotted separately by listener group, talker language background, and target predictability. Amplitude is plotted as a function of trial number. Individual points represent group means per trial; lines reflect model predicted values with shading reflecting standard error. YNH = younger normal hearing, ONH = older normal hearing.

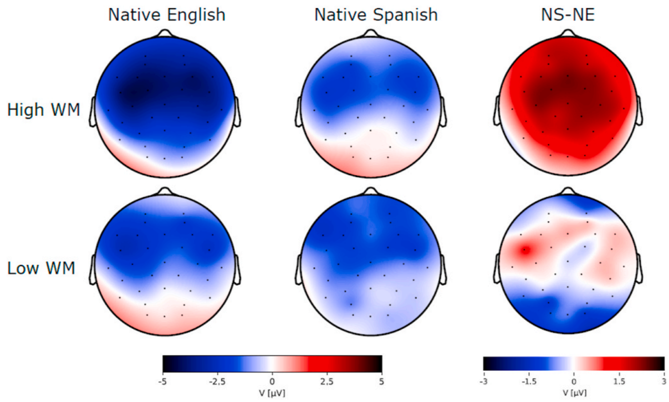

Fig. 12. Distribution of the response to the target word, averaged from 300 to 500 ms after word onset. The top row represents listeners with high working memory (WM) scores, and the bottom represents listeners with low WM scores. Responses to native English speech are shown on the left, and responses to Spanish-accented English are shown in the center. The right-hand panel shows the distribution of the talker effect (NS – NE).

measures. When examining the behavioral measures, there was a clear benefit of semantic context in the conditions including non-native speech stimuli. Both younger and older adults showed higher recognition scores and faster reaction times for HP sentences than LP sentences when listening to Spanish-accented English.

When examining the EEG measures, the response to the final target word of the sentence allows an examination of the influence of sentential context on lexical access for NE and NS speech. Both younger and older adults showed the expected N400 *effect of context*, which is interpreted to reflect a greater allocation of processing resources when speech is lacking in semantic context, due to the increased difficulty of lexical access for these items. Notably, this predictability effect did not differ across talkers, suggesting that, at this level of processing, the benefit of context was not impacted by talker accent. This stands in contrast to the behavioral findings, where the effect of context did differ by talker. Younger adults showed no context effect for native English speech in either the accuracy or reaction time measures, and older adults showed a relatively reduced context benefit in their accuracy scores.

One possible explanation for the discrepancies between behavioral and EEG measures is the fact that the current experimental paradigm does not allow us to ensure that the listeners were engaged in predictive processing – that is, listeners were instructed to report the final word of the sentence, and could have adopted a strategy of only listening for that word, in order to make their response.

The N400 effect can be elicited even by imperceptible stimuli (Deacon et al., 2000; Kiefer, 2002). That is, regardless of the listeners' strategy (attend to the whole sentence vs attend to the final word only), an N400 effect of context would be expected. If the listeners had disregarded the sentence frames, no effects of context would be expected on the behavioral outcome measures despite seeing evidence of the context manipulation in the EEG. Thus, it may be that the younger listeners adopted a final-word only strategy for the relatively 'easier' native English conditions, but made use of the sentence frames in the native Spanish conditions, rather relying on acoustics alone.

4.2. Effects of talker on acoustic processing

The EEG measures in this study allow for examination of several points in the speech recognition process. The earliest stage of auditory processing examined in this study, auditory object formation (as indexed by the P200 component), was impacted by non-native accent for older but not younger listeners. Older adults showed a reduction in the amplitude of the P200 response for the NS talker compared to the NE talker, suggesting that the older adults had more difficulty with auditory object formation for the NS as compared to the NE talker. This was seen despite the overall level of overcompensation shown in the older adults' P200 response (i.e. exaggerated P200 amplitude) relative to the younger listeners. The P200 component was also delayed in latency and increased in amplitude for older adults. These findings of delayed latency and exaggerated amplitude for the P200 are expected given prior literature about aging effects in the auditory cortex (Roque et al., 2019; Tremblay et al., 2003), and could contribute to the interactions with aging seen in the behavioral measures.

The effect of talker on P200 amplitude remained constant over the course of the experiment, suggesting that any learning of talker or accent was not evident in changes to the P200 response within this experiment. The stability of the P200, indicating no change to the processing of acoustic input, is consistent with prior work which has shown that training-induced changes to P200 do not occur within-session, but may emerge across sessions (Atienza et al., 2002; Rossi et al., 2013; Tremblay et al., 2001).

In response to the target word, both younger and older adults showed an N400 *effect of talker*, where the response to the NE talker had a greater negativity than the response to the NS talker. This effect was expected to occur in reverse: a greater negativity for the NS talker would have been interpreted as an increased difficulty of lexical access for NS speech, as was seen in response to the first word of the sentence. However, this reverse effect may have occurred as a side effect of the baselining procedure employed in the analysis. As is evident in the response to the first word of the sentence, the NS speech evoked an overall more negative response and a greater draw on processing resources at the outset of the sentence. This negativity was sustained over the duration of the sentence; see Fig. 13 for a visualization of the response to whole sentence at Cz.

4.3. Effects of talker and context on lexical processing

The first word of the sentence provides information about the ease of lexical access for the two talkers included in this study, before any contextual information becomes available. A centrally-distributed main effect of talker occurred just after 400 ms, consistent with an N400 effect. This finding is interpreted to reflect an increased difficulty in mapping lexical meaning onto NS speech, independent of any manipulations of predictability, and is consistent with the N400 effect of talker seen in prior literature involving non-native speakers (Goslin et al., 2012; Romero-Rivas et al., 2015). This finding is consistent with that seen by Romero-Rivas et al. (2015), who found that young adult listeners had larger N400 amplitudes in response to words spoken by non-native talkers as opposed to native talkers, independent of any context or predictability manipulations.

The main effect of talker remained constant over the course of the

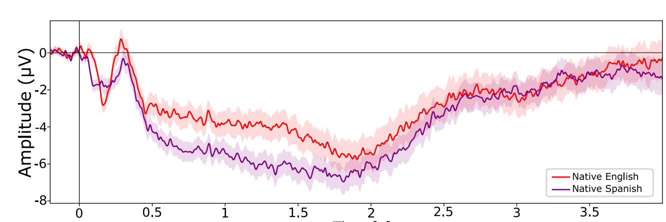

Fig. 13. Response to the entire sentence from the Cz electrode, separated by talker, showing a persistent negativity in response to Spanish-accented speech. The 0 ms timepoint represents the sentence onset. The target word onset occurred between 1.3 and 1.8 ms. Note that the Spanish-accented English speech evokes an overall more negative response for the first ~1.75 s. Thus, when the target word was heard (between ~1.3 and 1.8 s), this talker effect may reflect the fact that listeners had sparser processing resources available for the NE speech than the NS speech, allowing for greater magnitude of deflection in response to the NE target than the NS target.

experiment. The stability of the N400 response to the first word suggests that the process of lexical access required greater processing resources for the NS than for the NE talker, even with additional listening experience. This is unlike the findings of Romero-Rivas et al. (2015), who found a reduction in the magnitude of an N400 effect of talker across experimental blocks. One possible explanation for this discrepancy is that the prior study utilized eight accented talkers from different language backgrounds, while the present study used just one NS talker. Speech recognition is more challenging when listening to multiple talkers vs a single talker; perhaps the N400 effect of talker in the prior study was initially larger in magnitude than that of the present study, allowing more room for reduction over time.

4.4. Age, talker language, and predictability

While some interactions with age were observed in the behavioral response, the average EEG response to the target word did not show any age effects, nor interactions with age. This lack of any age effects contrasts with prior findings of an age-related reduction in the N400 effect (Federmeier et al., 2010; Federmeier and Kutas, 2005; Payne and Federmeier, 2018; Wlotko et al., 2012). Several factors could contribute to this finding. One potential discrepancy between this study and the prior ERP literature relates to hearing thresholds. Many of the ERP studies documenting an aging detriment did not measure or report pure-tone thresholds for their listener groups, which creates a potential confound between effects due to aging alone vs age-related hearing loss. Age-related hearing loss is known to result in detriments in auditory processing above and beyond age effects alone (Anderson et al., 2013a; Tremblay et al., 2003). In addition, the paradigm used in this study may have resulted in an increased degree of on-line attention required compared to tasks employed in prior literature. In this study, participants were asked to complete a word recognition task following each sentence, whereas previous studies employed tasks such as passive listening (Romero-Rivas et al., 2015), congruency judgements (Federmeier et al., 2003), or delayed recall (Federmeier and Kutas, 2005). The on-line word recognition task may have caused participants to devote relatively higher attention to the stimuli in this experiment, contributing to an improved ease of processing for the degraded signal (Wild et al., 2012) and eliminating the expected interaction.

Thus, effects of predictability, age, and talker accent emerged in various cortical measures of the speech perception process, and are observed in conjunction with an overall three-way-interaction present in the behavioral word identification scores. The effect of age on word identification accuracy may result more from age-related differences in the earlier processing of acoustic features and/or differences in listener strategy (see discussion above), as the higher-level predictive processing capabilities that are indexed by the context-elicited N400 do not seem to be reduced in the older normal-hearing adults, on average.

All listeners showed consistent improvements in word identification accuracy with additional trials. However, age effects did emerge when considering changes to N400 component magnitude and relative reaction times over the course of the experiment. Younger adults showed a steady decrease in N400 component magnitude with additional listening time for most conditions, while older adults did not. One potential explanation for this change is that the younger adults became globally more efficient at the process of lexical access across the course of the experiment and required fewer processing resources over time, while older adults did not. An alternate explanation could be that the younger adults experienced an increasing level of fatigue across the course of trials; this may also explain the increases in relative reaction time across conditions. The reaction time data also show that the older adults' relative reaction times decreased over the course of trials for all conditions, suggesting that their behavioral responses generally became less effortful over time.

Together, these results suggest that, while older adults may be able to improve their behavioral performance within the course of the experiment, the improvements likely do not stem from an increased ability to utilize predictive processing (N400 effect to target word) with additional listening experience. In addition, the older adults do not show evidence of an increased ability to extract acoustic or lexical information from the non-native speech input, independent of contextual manipulations (EEG responses to first word). Of course, some caution should be applied when directly comparing these data, as the cortical responses to the target word are only analyzed for trials where the behavioral response was correct, and the word identification accuracy measure considers all trials in each condition. Thus, the aim is not to directly correlate the two forms of measures, but rather to examine both commonly-used outcome measures in the same subjects to understand the patterns displayed by older and younger listeners.

One possible factor contributing to the older adults' improvements in behavioral performance over the course of the experiment could be age-related differences in the time course of task familiarization. In the current protocol, all listeners completed a practice round before beginning the experiment. However, this practice round only consisted of 8 trials, which may not have been sufficient for the older adult listeners to completely acclimate to the task, contributing to overall lower starting performance.

4.5. Working memory

Working memory capacity was found to predict N400 component amplitude in response to the target word of the sentence, with a stronger relationship between LSWMT scores and N400 amplitude for the NE speech. Working memory was not found to be predictive of the other ERP measures, nor the behavioral word identification measures. This finding extends prior literature documenting a relationship between working memory and a context-elicited N400 effect in response to visually presented stimuli (Federmeier and Kutas, 2005; Van Petten et al., 1997). Individuals with higher working memory scores showed greater N400 component amplitudes in response to NE speech. One possible explanation for this relationship is that these listeners had a greater ability to retain and manipulate the information contained in the sentences leading up to the target word, and thus greater processing resources were utilized in mapping meaning to the target word for these listeners. The absence of this relationship for the NS speech may relate to the overall increased demand imposed by processing more challenging speech. If listeners operate within a finite processing capacity, listeners may not have been able to draw on working memory resources in order to aid in the processing of the non-native speech.

Despite the relationship between LSWMT scores and N400 component amplitude, LSWMT scores did not emerge as a significant predictor of the behavioral responses. This was unexpected, given prior documentation of working memory as a significant predictor of speech recognition performance (Akeroyd, 2008; Anderson et al., 2013b; Banks et al., 2015; Füllgrabe et al., 2015; Lev-Ari, 2014). Indeed, working memory capacity is established as playing a critical role in speech understanding, especially in challenging listening environments (Füllgrabe et al., 2015). However, the lack of a significant relationship in the current study is not interpreted to indicate that working memory capacity does not contribute to speech recognition. Rather, we interpret these findings to suggest that the present measure of working memory was not sensitive to variation in word identification performance. N400 component amplitude reflects a particular aspect of language processing; it is possible that the working memory measure was sensitive to changes in this specific indicator of processing, and not behavioral keyword identification, which is a more gross measure. Another factor which potentially contributed to null effects such as the non-significant relationship between LSWMT scores and N400 component amplitude is the relatively small sample size included in this study. Future studies including a larger set of participants may be well suited to examine whether the contributions of working memory to speech understanding extend to a relationship between working memory measures and electrophysiologic measures of lexical access.

5. Conclusion

In conclusion, this study found that lexical access for speech produced by a non-native talker required greater processing resources than for speech produced by a native talker, regardless of listener age. Older adults did not show reductions in their ability to use context for lexical processing, as indexed by the predictability-elicited N400 effect, and talker native language did not influence the magnitude of this predictability effect. However, younger adults appeared to show increased ease of lexical access for both NE and NS speech with additional listening experience, while older adults did not. In contrast, older adults showed decreases in relative reaction time over the course of the experiment, while younger adults did not. However, all listeners showed improvements in word identification over time, which did not differ across talker or predictability conditions, despite overall poorer performance with the LP NS speech. Together, these findings expand the prior literature regarding aging and use of context in speech recognition, and suggest that within-session improvements in behavioral measures of word identification in older adults do not appear to result directly from improvements in predictive processing.

Credit author statement

Rebecca E. Bieber: Conceptualization, Investigation, Formal analysis, Writing – original draft, Writing – review & editing, Visualization. **Christian Brodbeck**: Conceptualization, Methodology, Software, Formal analysis, Writing – review & editing. **Samira Anderson**: Conceptualization, Resources, Writing – review & editing, Supervision.

Funding

This work was supported in part by training grant DC-00046 from the National Institute on Deafness and Communicative Disorders of the National Institutes of Health (Bieber), MCM Fund for Student Research Excellence (Bieber), NIH/NIDCD grant R21 DC015843 (Anderson).

Acknowledgements

The authors would like to thank Veronica Yevsukov for her assistance with data collection.

References

Adank, P., Janse, E., 2010. Comprehension of a novel accent by young and older listeners. Psychol. Aging 25, 736–740. https://doi.org/10.1037/a0020054.
Akeroyd, M.A., 2008. Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. Int. J. Audiol. 47, S53–S71. https://doi.org/10.1080/14992020802301142.
Anderson, S., Roque, L., Gaskins, C.R., Gordon-Salant, S., Goupell, M.J., 2020. Age-related compensation mechanism revealed in the cortical representation of degraded speech. JARO J. Assoc. Res. Otolaryngol. 21, 373–391. https://doi.org/10.1007/s10162-020-00753-4.
Anderson, S., White-Schwoch, T., Parbery-Clark, A., Kraus, N., 2013. A dynamic auditory-cognitive system supports speech-in-noise perception in older adults. Hear. Res. 300, 18–32. https://doi.org/10.1016/j.heares.2013.03.006.
Atagi, E., Bent, T., 2011. Perceptual Dimensions of Nonnative Speech. The 17th International Congress of Phonetic Sciences (ICPhS XVII).
Atienza, M., Cantero, J.L., Dominguez-Marin, E., 2002. The time course of neural changes underlying auditory perceptual learning. Learn. Mem. 9, 138–150. https://doi.org/10.1101/lm.46502.
Aydelott, J., Dick, F., Mills, D.L., 2006. Effects of acoustic distortion and semantic context on event-related potentials to spoken words. Psychophysiology 43, 454–464. https://doi.org/10.1111/j.1469-8986.2006.00448.x.
Banai, K., Lavie, L., 2020. Rapid perceptual learning and individual differences in speech perception: the good, the bad, and the sad. Audit. Percept. Cognit. 3, 201–211. https://doi.org/10.1080/25742442.2021.1909400.
Banks, B., Gowen, E., Munro, K.J., Adank, P., 2015. Cognitive predictors of perceptual adaptation to accented speech. J. Acoust. Soc. Am. 137, 2015–2024. https://doi.org/10.1121/1.4916265.
Bieber, R.E., Gordon-Salant, S., 2017. Adaptation to novel foreign-accented speech and retention of benefit following training: influence of aging and hearing loss. J. Acoust. Soc. Am. 141, 2800–2811. https://doi.org/10.1121/1.4980063.
Bilger, R.C., Nuetzel, J.M., Rabinowitz, W.M., Rzeczkowski, C., 1984. Standardization of a test of speech perception in noise. J. Speech Hear. Res. 27, 32–48. https://doi.org/10.1044/jshr.2701.32.
Brodbeck, C., Brooks, T.L., Das, P., Reddigari, S., Kulasingham, J.P., 2021. Eelbrain 0, p. 35. https://doi.org/10.5281/ZENODO.4650416.
Deacon, D., Hewitt, S., Yang, C.M., Nagata, M., 2000. Event-related potential indices of semantic priming using masked and unmasked words: evidence that the N400 does not reflect a post-lexical process. Cognit. Brain Res. 9, 137–146. https://doi.org/10.1016/S0926-6410(99)00050-6.
Dey, A., Sommers, M.S., 2015. Age-related differences in inhibitory control predict audiovisual speech perception. Psychol. Aging 30, 634–646. https://doi.org/10.1037/pag0000033.
Dubno, J.R., Ahlstrom, J.B., Horwitz, A.R., 2000. Use of context by young and aged adults with normal hearing. J. Acoust. Soc. Am. 107, 538–546. https://doi.org/10.1121/1.428322.
Federmeier, K.D., Kutas, M., 1999. A Rose by Any Other Name: Long-Term Memory Structure and Sentence Processing.
Federmeier, K.D., Kutas, M., 2005. Aging in context: age-related changes in context use during language comprehension. Psychophysiology 42, 133–141. https://doi.org/10.1111/j.1469-8986.2005.00274.x.
Federmeier, K.D., Kutas, M., Schul, R., 2010. Age-related and individual differences in the use of prediction during language comprehension. Brain Lang. 115, 149–161. https://doi.org/10.1016/j.bandl.2010.07.006.
Federmeier, K.D., McLennan, D.B., de Ochoa, E., Kutas, M., 2002. The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: an ERP study. Psychophysiology 39, 133–146. https://doi.org/10.1017/S0048577202001373.
Federmeier, K.D., Schwartz, T.J., Van Petten, C., Kutas, M., 2003. Sounds, words, sentences: age-related changes across levels of language processing. Psychol. Aging 18, 858–872. https://doi.org/10.1037/0882-7974.18.4.858.
Füllgrabe, C., Moore, B.C.J., Stone, M.A., 2015. Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition. Front. Aging Neurosci. 7, 347. https://doi.org/10.3389/fnagi.2014.00347.
Goslin, J., Duffy, H., Floccia, C., 2012. An ERP investigation of regional and foreign accent processing. Brain Lang. 122 (2), 92–102. https://doi.org/10.1016/j.bandl.2012.04.017.
Goy, H., Pelletier, M., Coletta, M., Pichora-Fuller, M.K., 2013. The effects of semantic context and the type and amount of acoustic distortion on lexical decision by younger and older adults. J. Speech Lang. Hear. Res. 56, 1715–1732. https://doi.org/10.1044/1092-4388(2013/12-0053.
Gramfort, A., Luessi, M., Larson, E., Engemann, D.A., Strohmeier, D., Brodbeck, C., Parkkonen, L., et al., 2014. MNE software for processing MEG and EEG data. Neuroimage 86, 446–460. https://doi.org/10.1016/j.neuroimage.2013.10.027.
Grey, S., van Hell, J.G., 2017. Foreign-accented speaker identity affects neural correlates of language comprehension. J. Neurolinguistics 42, 93–108. https://doi.org/10.1016/j.jneuroling.2016.12.001.
Hanulíková, A., van Alphen, P.M., Goch, M.M., Weber, A., 2012. When one person's mistake is another's standard usage: the effect of foreign accent on syntactic processing. J. Cognit. Neurosci. 24, 878–887. https://doi.org/10.1162/jocn_a_00103.
Hanulíková, A., Weber, A., 2012. Sink positive: linguistic experience with th substitutions influences nonnative word recognition. Atten. Percept. Psychophys. 74, 613–629. https://doi.org/10.3758/s13414-011-0259-7.
Hartman, M., Hasher, L., 1991. Aging and suppression: memory for previously relevant information. Psychol. Aging 6, 587–594.
Hox, J.J., Moerbeek, M., van de Schoot, R., 2010. Multilevel Analysis: Techniques and Applications, second ed. Routledge, New York.
Janse, E., 2012. A non-auditory measure of interference predicts distraction by competing speech in older adults. Aging Neuropsychol. Cognit. 19, 741–758. https://doi.org/10.1080/13825585.2011.652590.
Karawani, H., Jenkins, K., Anderson, S., 2018. Restoration of sensory input may improve cognitive and neural function. Neuropsychologia 114, 203–213. https://doi.org/10.1016/j.neuropsychologia.2018.04.041.
Kiefer, M., 2002. The N400 is modulated by unconsciously perceived masked words: further evidence for an automatic spreading activation account of N400 priming effects. Cognit. Brain Res. 13, 27–39. https://doi.org/10.1016/S0926-6410(01)00085-4.
Kutas, M., Hillyard, S.A., 1980. Reading senseless sentences: brain potentials reflect semantic incongruity. Science 207, 203–205. https://doi.org/10.1126/science.7350657.
Kutas, M., Hillyard, S.A., 1984. Brain potentials during reading reflect word expectancy and semantic association. Nature 307, 161–163.
Lau, E., Almeida, D., Hines, P.C., Poeppel, D., 2009. A lexical basis for N400 context effects: evidence from MEG. Brain Lang 111, 161–172. https://doi.org/10.1016/j.bandl.2009.08.007.
Lau, E.F., Phillips, C., Poeppel, D., 2008. A cortical network for semantics: (De)constructing the N400. Nat. Rev. Neurosci. https://doi.org/10.1038/nrn2532.
Lev-Ari, S., 2014. Comprehending non-native speakers: theory and evidence for adjustment in manner of processing. Front. Psychol. 5, 1–12. https://doi.org/10.3389/fpsyg.2014.01546.

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

Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. J. Neurosci. Methods 164, 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024.

Mirman, D., 2014. Growth curve analysis: A hands-on tutorial on using multilevel regression to analyze time course data. Proceedings of the Annual Meeting of the Cognitive Science Society 36 (36).

Nasreddine, Z.S., Phillips, N.A., Bã©dirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J.L., et al., 2005. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J. Am. Geriatr. Soc. 53, 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x.

Payne, B.R., Federmeier, K.D., 2018. Contextual constraints on lexico-semantic processing in aging: Evidence from single-word event-related brain potentials. Brain Res. 1687, 117–128. https://doi.org/10.1016/j.brainres.2018.02.021.

Pichora-Fuller, M.K., 2008. Use of supportive context by younger and older adult listeners: balancing bottom-up and top- down information processing. J. Int. J. Audiol. 47, 1499–2027. https://doi.org/10.1080/14992020802307404.

Pichora-Fuller, M.K., Schneider, B.A., Daneman, M., 1995. How young and old adults listen to and remember speech in noise. J. Acoust. Soc. Am. 97, 593–608. https://doi.org/10.1121/1.412282.

Pichora-Fuller, M.K., Schneider, B.A., Daneman, M., 1995a. How young and old adults listen to and remember speech in noise. J. Acoust. Soc. Am. 97, 593. https://doi.org/10.1121/1.412282.

Pichora-Fuller, M.K., Schneider, B.A., Daneman, M., Pichora-Fuller, M.K., Schneider, B. A., Daneman, M., 1995b. How young and old adults listen to and remember speech in noise. J. Acoust. Soc. Am. 97, 593–608. https://doi.org/10.1121/1.412282.

Reinke, K.S., He, Y., Wang, C., Alain, C., 2003. Perceptual learning modulates sensory evoked response during vowel segregation. Cognit. Brain Res. 17, 781–791. https://doi.org/10.1016/S0926-6410(03)00202-7.

Rogers, C.S., Wingfield, A., 2015. Stimulus-independent semantic bias misdirects word recognition in older adults. J. Acoust. Soc. Am. 138, EL26. https://doi.org/10.1121/1.4922363. –EL30.

Romero-Rivas, C., Martin, C.D., Costa, A., 2015. Processing changes when listening to foreign-accented speech. https://doi.org/10.3389/fnhum.2015.00167.

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., Dahlström, Ö., et al., 2013. The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. Front. Syst. Neurosci. 7, 1–17. https://doi.org/10.3389/fnsys.2013.00031.

Rönnberg, J., Rudner, M., Foo, C., Lunner, T., 2008. Cognition counts: a working memory system for ease of language understanding (ELU). Int. J. Audiol. 47, S99–S105. https://doi.org/10.1080/14992020802301167.

Roque, L., Karawani, H., Gordon-Salant, S., Anderson, S., 2019. Effects of age, cognition, and neural encoding on the perception of temporal speech cues. Front. Neurosci. 13, 749. https://doi.org/10.3389/fnins.2019.00749.

Rossi, S., Hartmüller, T., Vignotto, M., Obrig, H., 2013. Electrophysiological evidence for modulation of lexical processing after repetitive exposure to foreign phonotactic rules. Brain Lang. 127 (3), 404–414. https://doi.org/10.1016/j.bandl.2013.02.009.

Schurman, J., Brungart, D., Gordon-Salant, S., 2014. Effects of masker type, sentence context, and listener age on speech recognition performance in 1-back listening tasks. J. Acoust. Soc. Am. 136, 3337–3349. https://doi.org/10.1121/1.4901708.

Sheldon, S., Pichora-Fuller, M.K., Schneider, B., 2008. Priming and sentence context support listening to noise-vocoded speech by younger and older listeners. J. Acoust. Soc. Am. https://doi.org/10.1121/1.2783762.

Sommers, M.S., Danielson, S.M., 1999. Inhibitory processes and spoken word recognition in young and older adults: the interaction of lexical competition and semantic context. Psychol. Aging 14, 458–472. https://doi.org/10.1037/0882-7974.14.3.458.

Straus, A., Kotz, S., Obleser, J., 2013. Narrowed expectancies under degraded speech: revisiting the n400. J. Cognit. Neurosci. 26, 194–198. https://doi.org/10.1162/jocn.

Tremblay, K., Kraus, N., McGee, T., Ponton, C., Otis, A.B., 2001. Central auditory plasticity: changes in the N1-P2 complex after speech-sound training. Ear Hear. 22, 79–90. https://doi.org/10.1097/00003446-200104000-00001.

Tremblay, K.L., Piskosz, M., Souza, P., 2003. Effects of age and age-related hearing loss on the neural representation of speech cues. Clin. Neurophysiol. 114 (7), 1332–1343.

Van Petten, C., Weckerly, J., McIsaac, H.K., Kutas, M., 1997. Working memory capacity dissociates lexical and sentential context effects. Psychol. Sci. 8 (3), 238–242.

Weintraub, S., Dikmen, S.S., Heaton, R.K., Tulsky, D.S., Zelazo, P.D., Bauer, P.J., Carlozzi, N.E., et al., 2013. Cognition assessment using the NIH Toolbox. Neurology 80, S54–S64. https://doi.org/10.1212/WNL.0b013e3182872ded.

Wild, C.J., Yusuf, A., Wilson, D.E., Peelle, J.E., Davis, M.H., Johnsrude, I.S., 2012. Effortful listening: the processing of degraded speech depends critically on attention. J. Neurosci. 32 (40), 14010–14021. https://doi.org/10.1523/JNEUROSCI.1528-12.2012.

Wingfield, A., Alexander, A.H., Cavigelli, S., 1994. Does memory constrain utilization of top-down information in spoken word recognition? Evidence from normal aging. Lang. Speech 37 (Pt 3), 221–235. https://doi.org/10.1177/002383099403700301.

Wlotko, E.W., Federmeier, K.D., Kutas, M., 2012. To predict or not to predict: age-related differences in the use of sentential context. Psychol. Aging 27, 975–988. https://doi.org/10.1037/a0029206.