JULY 16 2025

Context effects for degraded speech: Effects of age, preceding or subsequent contextual cues, and signal-to-noise ratio

Anna R. Tinnemore ; Sandra Gordon-Salant ; Matthew J. Goupell

J. Acoust. Soc. Am. 158, 447–460 (2025)
https://doi.org/10.1121/10.0037182

JASA ARTICLE

Context effects for degraded speech: Effects of age, preceding or subsequent contextual cues, and signal-to-noise ratio

Anna R. Tinnemore,[1,2,a)] Sandra Gordon-Salant,[1,2] and Matthew J. Goupell[1,2]
[1]Neuroscience and Cognitive Science Program, University of Maryland, College Park, Maryland 20742, USA
[2]Department of Hearing and Speech Sciences, University of Maryland, College Park, Maryland 20742, USA

ABSTRACT:
Predictive sentence contexts can be used to support speech understanding when words are degraded or unclear. Older adults are thought to maintain the ability to benefit from context. Because context effects are usually measured on words at the end of a sentence, it is unknown if a word's location in a sentence interacts with a listener's age and the word's degradation level to influence the context effect. In this study, listeners (20–76 years) with normal hearing were presented spectrally degraded (8-channel vocoded) speech and performed a phoneme categorization task for words embedded in various levels of speech-shaped noise at the beginning or end of sentences with congruent, incongruent, or neutral contexts. Phoneme categorization accuracy and response times were measured. Results showed effects of target word location within the sentence, especially at more difficult signal-to-noise ratios (–5 and −10 dB). Although there was no significant effect of age on the magnitude of the context effect, there were significant interactions between age, signal-to-noise ratio, and sentence position on response times. These findings suggest that listeners' context benefit depends on a degraded word's position within the sentence and support the theory that all listeners, including older adults, can benefit from context cues.
https://doi.org/10.1121/10.0037182

(Received 6 December 2024; revised 15 May 2025; accepted 22 June 2025; published online 16 July 2025)

[Editor: Christopher A. Brown] Pages: 447–460

I. INTRODUCTION

Successful communication involves decoding sensory signals that carry messages between people. This decryption of sensory input is facilitated by combining bottom-up information (e.g., peripheral inputs and central auditory processing) with top-down information (e.g., cognitive skills, resources, and effort) (Marslen-Wilson and Tyler, 1980; Morton, 1969). A listener's knowledge of vocabulary, grammar, and the rules governing semantic roles is combined with the available sensory information, even if degraded, to narrow the range of possible intended messages and provide a context benefit (e.g., Fletcher *et al.*, 2019; McClelland and Elman, 1986). This knowledge of the language can be used to overcome sensory deficits—simulated in the current study with the use of spectrally degraded speech and various levels of background noise—through the top-down use of context (Amichetti *et al.*, 2018; Sohoglu *et al.*, 2012; Sommers and Danielson, 1999; Wild *et al.*, 2012).

Context is defined as the information surrounding a target degraded word or sound in space and time. This contextual information can be acoustic, phonemic, lexical, semantic, syntactic, or a combination of the above. Thus, a "context effect" is any change in the perception of a word or sound in the presence of contextual information compared with the perception of the same word or sound in the absence of contextual information. Context effects on identifying sentence-final words or on recognizing phonemes within a word are most evident when the target is degraded (e.g., Bhandari *et al.*, 2021; Boothroyd and Nittrouer, 1988). Speech can be degraded in multiple ways; perhaps the most common of which is the presence of background noise. Non-relevant background noise can mask the spectral and temporal acoustic features of the target speech signal, making the target difficult or even impossible to understand.

Increasing age is often accompanied by declines in sensory acuity (e.g., hearing loss) and potential declines in cognitive abilities (Cruickshanks *et al.*, 1998; Gordon-Salant *et al.*, 2010; Humes *et al.*, 2013; Lin, 2011; Lipnicki *et al.*, 2017; Park *et al.*, 2002). At the same time, increasing age is also associated with increased vocabulary size and extensive language experience (Ben-David *et al.*, 2015; Kavé and Halamish, 2015; Milburn *et al.*, 2023; Verhaeghen, 2003). Context usage, bolstered by language experience and vocabulary, may be sufficient to offset the declines in both sensory and cognitive abilities that often occur with increased age, allowing older adults to match younger adults' high levels of speech recognition performance.

There are many methods that can be used to measure context effects. Some involve open-set word identification tasks for target words at the end of a sentence, such as the Revised Speech In Noise test (Bilger *et al.*, 1984), while others involve closed-set word or phoneme recognition, such as a phoneme categorization task with target words embedded in a sentence (e.g., Connine, 1987; Connine and

a)Email: annat@umd.edu

Clifton, 1987). Because speech can be described as hierarchical (e.g., Ding *et al.*, 2016; Jurafsky, 1996; McClelland and Elman, 1986) and any disruption of understanding at low levels of the speech hierarchy could impact understanding of the whole sentence, a phoneme categorization task was chosen to assess the impact of contextual information on perception at a low level of the speech hierarchy (e.g., Connine and Clifton, 1987; Liberman *et al.*, 1967; Liberman *et al.*, 1977; Repp *et al.*, 1984). In this task, a participant hears words that vary along a continuum on a single acoustic dimension between two distinct endpoints (e.g., dent and tent). The participant must categorize each stimulus as either one endpoint or the other.

A. Factors affecting the magnitude of the context effect

The relative locations of a target word and contextual information within a sentence may influence the magnitude of a potential context effect, across multiple task types. The few studies that have assessed a context effect on a target word that was followed by contextual information rather than preceded by contextual information have found smaller, but still measurable, context effects compared to context effects on sentence-final target words (e.g., Connine *et al.*, 1991; Wingfield *et al.*, 1994; Wotton *et al.*, 2011). A context effect, as measured by a difference in phoneme categorization by young listeners, was observed when contextual information occurred within three syllables after the target word (Connine *et al.*, 1991). The context effect disappeared when the contextual information occurred later in the sentence. That same study showed a context effect on the response times for categorizing ambiguous target words. Another study showed context effects on the recognition of reverberant vowels within words embedded near the beginning of congruent, incongruent, or neutral sentences, even when the contextual information occurred more than three syllables from the target word (Wotton *et al.*, 2011). In another study employing an open-set word recognition task (Wingfield *et al.*, 1994), younger and older listeners showed a relatively small context benefit from context following target words compared to context preceding target words. In addition, when the context followed the target words, older listeners showed significantly less benefit from that context than younger listeners. Thus, the target word's position in the sentence and the age of the listener appear to interact to affect open-set word recognition.

Signal degradation can also affect various speech recognition tasks, including phoneme categorization, the task used in the current study. Spectral degradation and background noise are two common types of signal degradation. Young listeners (18–31 years) with normal hearing who are presented spectrally degraded isolated words show poorer phoneme categorization with decreasing levels of spectral resolution (Winn and Litovsky, 2015). Context benefit measured in other types of speech identification tasks is greatest when the sensory input is moderately degraded. This has been shown with spectral degradation, such as cochlear-implant processed speech (e.g., Bhandari *et al.*, 2021; Obleser *et al.*, 2007; Obleser and Kotz, 2011; Sohoglu *et al.*, 2012; Wild *et al.*, 2012), and with background noise (e.g., Golestani *et al.*, 2009; Miller *et al.*, 1951).

Both sentence context and signal degradation affect speech recognition, including when measured using a phoneme categorization task. However, it is unclear how a listener's age interacts with the amount of signal degradation and the target word's position in the sentence to influence the benefit of context. On the one hand, older adults, even those with normal hearing, exhibit considerable difficulty on speech recognition tasks in noise (e.g., Dubno *et al.*, 1984; Gordon-Salant and Cole, 2016; Helfer and Wilber, 1990). On the other hand, older adults exhibit good use of contextual information to support speech recognition in noise (e.g., Pichora-Fuller, 2008; Pichora-Fuller *et al.*, 1995; Sommers and Danielson, 1999). Therefore, one question addressed in this study is whether or not older listeners use contextual information differently than younger listeners for degraded words presented in varying sentence locations. To address this question, this study tested adults from a wide range of ages on a phoneme categorization task for contrasting phonemes in a single word pair (henceforth referred to as the phoneme categorization task in this paper). Spectrally degraded target words were presented at the beginning or end of spectrally degraded context sentences. The context sentences were presented in quiet to ensure audibility, and only the target words were presented in background noise. The context effect was calculated based on the differences in categorization between these conditions.

B. Experimental questions

There were three main research questions: (1) Does the position of a spectrally degraded and noise-masked target word in a spectrally degraded sentence alter the effects of context on phoneme categorization? (2) Is the age of the listener predictive of the extent to which listeners use context when identifying a noise-masked target word in a spectrally degraded sentence? (3) Is there an interaction between the effect of sentence position and age, such that the effect of sentence position is smaller for older listeners than younger listeners?

Given the previously measured small effects for subsequent context (Connine *et al.*, 1991; Wingfield *et al.*, 1994), the hypothesis for the first research question was that the context effect would be greater when the target word was at the end of the sentence than when it was at the beginning of the sentence. A noise-masked target word at the end of the sentence may be perceived as congruent with the predictive sentence because the degraded acoustics might not override the brain's prediction of the final target word (Federmeier, 2007; Hannemann *et al.*, 2007; Saija *et al.*, 2014). The identification of a noise-masked target word at the beginning of the sentence should be less affected by subsequent context because no internal prediction had been made prior to hearing the target word.

JASA https://doi.org/10.1121/10.0037182

The hypothesis for the second research question was that age would not be predictive of the context effect on categorization performance (given the findings of O'Neill *et al*., 2021) but would be predictive of the response times of participants. Response times are considered an objective measure of the decision-making process. It was hypothesized that increasing age would correspond with increasingly longer response times to target words occurring in incongruent sentence contexts. This slowed response could occur because of age-related reductions in processing speed and in the ability to inhibit the influence of incongruent information (e.g., Hasher *et al*., 1991).

The hypothesis for the third research question was that, with increasing age, participants would show smaller differences in the context effects between the two sentence positions. Similar effects of sentence context on categorization of target phonemes within words are expected across ages when the target word is at the end of the sentence, while smaller effects of sentence context are expected with increasing age when the target word is at the beginning of the sentence (Wingfield *et al*., 1994).

II. METHOD

A. Participants

Thirty-six adult participants (32 female, ages 20–76 years, mean = 48.7 years, SD = 18.7 years) with nominally normal acoustic hearing were recruited. All participants had no more than a slight hearing loss (thresholds ≤ 25 dB HL) at octave frequencies between 250 and 4000 Hz in at least one ear (re: ANSI, 2018) (Fig. 1). Thresholds were measured using supra-aural headphones (TDH-50P, Telephonics, Santa Ana, CA) on a clinical audiometer (Audiostar Pro, Grason-Stadler, Inc., Eden Prairie, MN). There was no statistically significant correlation between age and the four-frequency pure tone average of participants' hearing thresholds [500, 1000, 2000, and 4000 Hz; $r(34) = -0.15$, $p = 0.394$], indicating that average hearing sensitivity did not change significantly with increasing age. Participants were self-reported native speakers of American English to avoid variability on the speech identification tasks stemming from diverse language backgrounds. There were no significant differences in education as a function of age [$r(33) = -0.07$, $p = 0.689$]. Education level is sometimes used as a proxy for vocabulary knowledge (e.g., Verhaeghen, 2003). Participants had to be able to read moderately large print on a computer screen with or without corrective lenses, which was determined during the practice trials. Participants were primarily recruited via word of mouth and a database of participants who had previously consented to be contacted with future research opportunities.

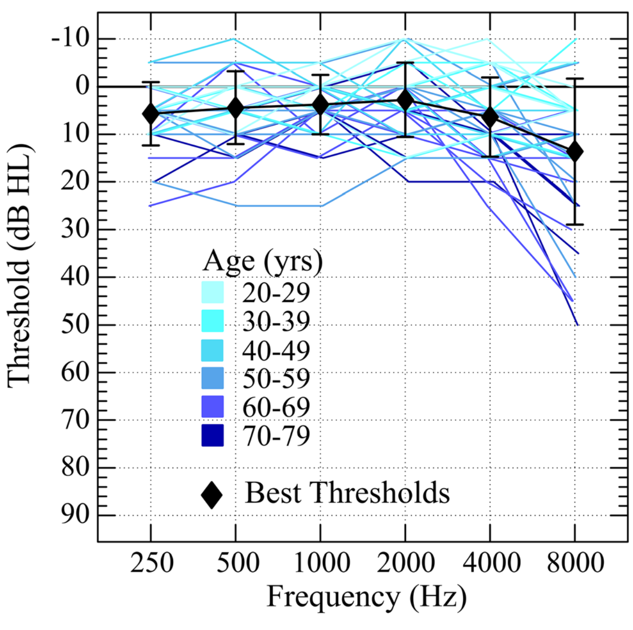

FIG. 1. Individual participants' best audiometric thresholds are shown at octave frequencies between 250 and 8000 Hz The color gradient of the lines represents the age of the participants (lighter = younger). The average thresholds for all participants are represented using black diamonds. Error bars indicate ±1 standard deviation.

B. Stimuli

The target stimuli were audio recordings of words presented at the beginning or end of sentences with or without predictive semantic contexts. Stimuli were created from natural target words produced by a 43-year-old male native speaker of American English (mean f0 = 125 Hz). Recordings took place in a double-walled sound-treated booth (IAC Acoustics, Naperville, IL). The talker was seated so that his mouth was positioned 8 inches away from the wind guard covering a condenser microphone (Shure KSM141, Niles, IL) set in omni-directional mode with zero gain and a flat frequency response. The microphone was connected to an audio recorder (Zoom H4n Pro, Hauppauge, NY). The recording sampling rate was 44 100 Hz with 16-bit depth. The talker recorded each stimulus word and sentence multiple times. After a recording session, the recordings were transferred to a computer for excision and analysis.

The context sentences chosen as the auditory stimuli for each condition were selected from the multiple recordings based on the researcher's judgment of clarity and intelligibility. The natural target words chosen for stimulus development were recorded at the end of the context sentences to retain the natural speech rhythm and prosody of the talker before being excised for processing. A four-step continuum was created by varying the voice onset time (VOT) from 26 to 86 ms for a word-initial /d/ to /t/, such that target words varied perceptually between the rhyming words, "deer" and "tear" (/diɹ/ and /tiɹ/). Each step along the continuum contained an additional 20 ms of the aspiration noise taken from the talker's natural production of "tear" added after the initial burst of the "deer" naturally produced word (Fig. 2). The endpoints of the continuum, Step 1 with 26 ms of VOT (total stimulus length = 310 ms) and Step 4 with 86 ms of VOT (total stimulus length = 370 ms), were clear tokens of the two target words, "deer" and "tear." The intermediate steps could be perceived as either of the two target words.

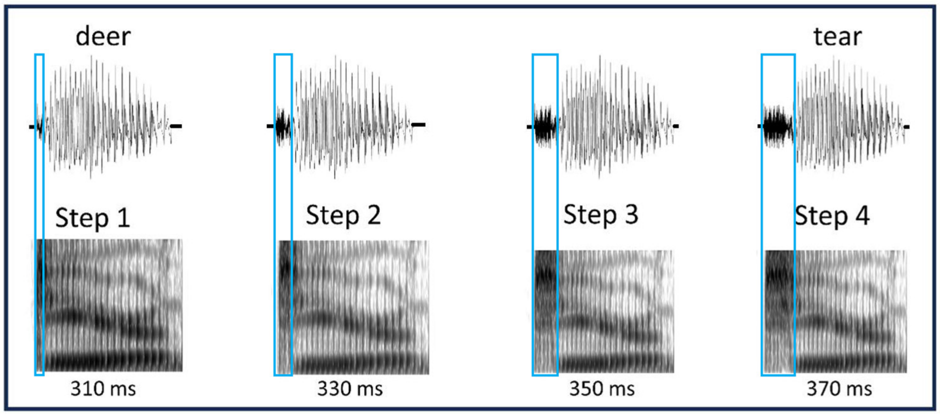

FIG. 2. Waveforms and spectrograms of the four steps along the dear-tear continuum. The boxes indicate the initial aspiration that was added to create the continuum. The duration of the word in milliseconds (ms) is provided under each step.

Each continuum step was concatenated to the article "the" so that the target words could be presented either before or after the context sentence in the presentation software. Between the article and the noun were 70 ms of silence, the average length of time between these two words in the intact recorded sentences. The context sentences and the target word audio files ("the dear" and "the tear") were equalized in root mean square amplitude. A single word pair was chosen so that the effects of context, sentence location, and signal degradation could be thoroughly explored.

In a pilot study, five young participants with normal hearing provided behavioral judgments that confirmed that participants perceived the stimuli as steps along a continuum. The two target words at the endpoints were randomly presented in isolation 5–10 times each. They were identified with 100% accuracy, confirming that the natural "deer" stimulus was perceived as "deer" and the created "tear" stimulus was easily perceived as "tear." Sentence stimuli were created that provided either neutral semantic context or contained at least two words that were semantically related to one of the words at the continuum endpoints ("deer" or "tear") (as in Kalikow *et al.*, 1977). The sentences used for context are presented in Table I. When pilot participants were presented the context sentences without an audible target word and asked whether "deer" or "tear" best completed the sentence, they always chose "deer" to complete the "deer" context sentences and "tear" to complete the "tear" context sentences.

All auditory stimuli (target words and context sentences) were spectrally degraded with a sine vocoder (Cychosz *et al.*, 2024). The auditory waveforms were divided into eight logarithmically spaced channels between 250 and 4000 Hz using third-order forward-backward Butterworth filters. Eight vocoded channels produce a signal with a moderate level of spectral degradation that is still highly intelligible for listeners with normal hearing (Dorman *et al.*, 1997; Friesen *et al.*, 2001). Envelopes were extracted by half-wave rectifying the contents of each channel and then using a 150-Hz second-order forward-backward Butterworth filter (Shader *et al.*, 2020). All stimuli were low-pass filtered at 4000 Hz to avoid potential confounds with high-frequency age-related hearing loss, as in Tinnemore *et al.* (2020). In addition, the VOT contrast in the deer/tear continuum does not critically rely on energy above 4000 Hz. The background noise was speech-shaped noise that was created from the long-term average spectrum of the target words and sentences used in this experiment. The noise was presented starting 400 ms before the target word to avoid overshoot (Bacon and Healy, 2000) and stopping at the end of the target word. There was no background noise during the context sentences. The background noise was added to the target stimuli at 10, 5, 0, −5, and −10 dB signal-to-noise ratio (SNR) before the entire signal was vocoded. This range of SNRs was chosen so that the degradation provided by the combination of vocoding and background noise would be likely to include each participant's unique "moderately degraded" level of background noise, known to cause maximum context effects (Bhandari *et al.*, 2021; Boothroyd and Nittrouer, 1988).

TABLE I. Sentences used to provide context for sentence-initial and sentence-final presentation of continuum words.

Sentence position	Context type	Sentence stimuli
Initial	Neutral	"The ____ in the picture was unclear."
Initial	Deer	"The ____ lost its antlers in the forest."
Initial	Tear	"The ____ fell from her eye down her cheek."
Final	Neutral	"The girl asked her mom about the ____."
Final	Deer	"The hunter saw the antlers of the ____."
Final	Tear	"The tissue caught the moisture of the ____."

C. Procedure

The experimental procedures were approved by the University of Maryland Institutional Review Board for Human Subjects Research. Participants provided written informed consent and were paid for their time. Participants were seated comfortably in a sound-attenuating booth. Auditory stimuli (context words and target sentences) were presented over a loudspeaker (Yamaha HS5, Buena Park, CA) approximately 1 m directly in front of the participant's chair for most of the participants or over circumaural headphones (Sennheiser HD-650, Old Lyme, CT) for seven participants, because of limitations in equipment availability. There were no statistically significant differences in context effect [$t(33) = -0.827$, $p = 0.414$] or response time

[$t(33) = 0.398$, $p = 0.693$] between the participants who listened over headphones and those who listened over the loudspeaker. The stimuli were calibrated in the sound field to be presented at 70 dB-A using a sound level meter (Brüel & Kjaer type 2250, Naerum, Denmark) with a quarter-inch microphone to measure a segment of speech-shaped noise set to the same root mean square amplitude as the target stimuli. A 1000-Hz tone with the same root mean square amplitude as the target stimuli was used with the same sound level meter connected to a 2-cm^3 coupler to calibrate the headphones that were used for seven of the participants.

1. Practice

The visual stimuli were presented on a 23-inch computer monitor located 0.75 m from the participant's chair. Prior to beginning the main experiment, participants completed a short practice session during which they heard the context sentences that would be used in the experiment. The practice session served many purposes, including the familiarization of the participant to the task and to the auditory and visual stimuli. During each self-initiated trial, the word "Ready…" showed on the screen for 1.2 s. The sound files of the context sentence and target word played while the screen showed the printed sentence they were hearing, with the target word represented by a blank (as in Table I). After the auditory presentation, the question "Which word did you hear?" appeared on the top half of the screen and two color images that represented the endpoint target words appeared on the bottom half of the screen to either side. At that point, the subroutine recording the participants' responses became active. No early responses were accepted. The mouse cursor became visible at a centered position equidistant from the two response buttons. Participants used the mouse to select the picture representing the chosen word on the computer monitor. Response times were measured from the time the mouse became visible (at the end of the sound file). Every trial had the same response options that were always in the same locations on the screen. During this practice session, only the endpoint words (Steps 1 and 4) were presented. Participants were informed that occasionally the word might not match the sentences.

The practice session was divided into four distinct blocks, each containing 12 trials. In the first block, the vocoded context sentences and target words were presented in quiet. Two of the trials contained target words that did not match the context sentence (i.e., incongruent contexts). Correct-answer feedback was provided. Trials were presented in a random order. In the second block, the context sentences were neutral (i.e., no predictive contexts) and the target words were presented in quiet. This block was used to ensure that participants were able to identify the spectrally degraded endpoint words with at least 90% accuracy and to collect a baseline measurement of the participant's response times. Although these individualized response time data were collected, they were not used in the analysis (see Sec. II C 4 for more details). In the third block, the target words were presented in background noise in neutral sentence contexts. Participants first heard three trials with the target words presented at 10 dB SNR. Then, two trials with target words at each of the remaining four SNRs were presented, becoming progressively more difficult. The last trial's target word was presented in quiet. In the final block of the practice, baseline response times to target words presented at the most favorable SNR (10 dB SNR) in neutral context sentences were collected. Again, these response times were not used in the current analysis. In summary, the practice session in its entirety accomplished several objectives. It ensured that the context sentences would be easily recognized and understood during the task. It allowed some additional adaptation to the vocoded speech. It provided familiarity with the task. Last, it confirmed that all participants could correctly identify the endpoint target words in both quiet and 10 dB SNR background noise with at least 90% accuracy.

2. Main experiment

During the main experiment, the screen was blank during the presentation of the context sentence and target word. After the auditory presentation was complete, the question "Which word did you hear?" and the pictures representing the appropriate endpoint words appeared together with the mouse cursor, as in the practice session. Participants used the mouse to select the picture representing their choice for the target word. The experiment had six conditions: the target words presented at the beginning (one condition) or end (second condition) of a sentence that did not provide predictive context (i.e., neutral context), the target words presented at the beginning (third condition) or end (fourth condition) of sentences that predicted "deer," and the target words presented at the beginning (fifth condition) or end (sixth condition) of sentences that predicted "tear."

Target words from each step along the continuum were presented 10 times at the beginning and end of each of the three types of sentence contexts at five different SNRs (4 target words from the phonetic continuum $\times$ 3 sentence contexts $\times$ 2 sentence positions $\times$ 10 repetitions $\times$ 5 SNRs $=$ 1200 trials). Each block contained one trial of each condition (120 trials). The presentation order was fully randomized within each block. An opportunity to take a break was provided after every block or approximately every 15 min to avoid fatigue. Performance was measured in two ways: (1) a percent "deer" response for each target word along the continuum; and (2) response times for categorization choices.

3. Cognitive tests

After the last block of the experiment, participants completed two cognitive measures: (1) verbal processing speed and inhibition (Stroop task; Stroop, 1935); and (2) working memory (NIH Toolbox List Sorting; Tulsky *et al.*, 2014). The age-normalized scores from these measurements were used in a trial analysis as predictors of the effect of context

https://doi.org/10.1121/10.0037182

on speech recognition performance. Neither measure was a significant predictor, and thus neither measure is reported in the results for clarity and simplicity.

Participants completed the experiment over two or three sessions. Each session was no more than 2 hours in duration and was scheduled on a different day. The average participant completed the practice, the main experiment, and the cognitive tests over the course of two sessions (∼4 hours). Some of the older participants required a short third session.

4. Analysis

Two analyses were conducted to assess the context effect on the perception of phonetic cues in two sentence positions and with varying levels of background noise. The analyses modeled the context effect as either behavior-based (derived from performance on a phoneme categorization task) or timing-based (response times during the phoneme categorization task). Specific analyses varied with each processing measure, as described below.

The context effect was calculated by comparing an individual's phoneme categorization performance in the neutral context condition with their performance in the predictive context conditions. A positive context effect would reflect more target words categorized as the context-predicted word compared with target word categorizations in the neutral context. In this experiment, a positive context effect could be a result of more "deer" responses in a deer-predictive context sentence or fewer "deer" responses in a tear-predictive context sentence compared with the neutral context. The comparison to the neutral context condition accounts for response biases stemming from the target word's position, SNR, and VOT, because the target word in the neutral context condition is categorized under the exact same position, SNR, and VOT conditions. Some, but not all, of the previous studies used a neutral comparison to determine their context effects. The current method is unique because it allows the calculated context effect to be negative. An example of a negative context effect would be more "deer" responses in the neutral context condition than in the deer-predictive context condition. Negative context effects do not make logical sense in real-world environments but could represent a listener strategy on our task. Allowing context effects to be negative, rather than a default value of zero if the difference was not positive, may minimize the calculation of the context effect. However, the inclusion of negative context effects strengthens confidence that a context effect is or is not statistically greater than zero.

The focus of this study is the overall effect of sentence position, SNR, and presence of predictive context on perception of a target word varying in its initial phoneme. Thus, the difference in categorization performance was calculated for each listener in each sentence position, SNR, VOT, and predictive context condition and then summed across VOTs and across predictive contexts. The context effect, as calculated, represented the entire change in categorization performance (for all VOT versions of a target word) that was associated with the presence of either predictive sentence context compared to the neutral sentence context as a function of sentence position and SNR. The statistical analysis was conducted in R (version 4.3.3; R Core Team, 2024) using the *lme4* package (version 1.1.35.3; Bates *et al.*, 2015).

The log-transformed context effect from each participant in each condition was modeled with a linear mixed-effect model (MEM) with a Gaussian distribution. The fixed effects were age, sentence position (initial or final), SNR (5 levels), and their interactions. Age and SNR were treated as continuous variables. Both variables were standardized such that the mean of each variable (mean age = 49.2 years, mean SNR = 0 dB SNR) was zero (with a standard deviation of 1) and was the reference value in the model. Sentence position was a categorical variable with a reference value of sentence-final position. The maximal random effects structure included random intercepts for each participant, with random slopes of SNR and sentence position (Barr *et al.*, 2013).

In summary, the analysis modeled the effects of age, sentence position, SNR, and their interactions on the derived context effects. The intercept of this model represented the context effect measured for a participant at the mean age (49.2 years) in the 0 dB SNR condition when the target word was presented at the end of the sentence. The effects of both SNR and sentence position were allowed to be different for each participant, and these effects were captured in the random effects.

In the second analysis, the absolute response times for all trials were modeled in a linear MEM (Gaussian distribution). Response times less than 0.1 s and greater than 10 s were removed from the analysis because response times less than 0.1 s are physically impossible after a decision, while response times greater than 10 s are likely to represent lapses in attention or other factors not related to decision making. This left 43 023 response times to analyze. A log-transformation was applied to each response time value (plus a constant value of 10 to eliminate negative numbers) to reduce the skewness of the distribution and approximate a normal distribution. No further normalization was needed because the MEM accounts for individual baseline response times in the random effect of participant. Therefore, the response times collected during the practice sessions were not needed for the analysis.

The fixed effects were SNR, target word acoustical ambiguity, age, sentence position, and context sentence congruency. SNR and age were treated as continuous variables. Both variables were standardized such that the mean of each variable (mean age = 49.2 years, mean SNR = 0 dB SNR) was zero (with a standard deviation of one) and was the reference value in the model. A categorical variable to represent acoustic ambiguity was created using the steps along the VOT continuum. The two endpoint steps were categorized as acoustically unambiguous, while the two middle steps were categorized as acoustically ambiguous. The unambiguous category was used as the reference. Sentence

https://doi.org/10.1121/10.0037182

position was a categorical variable with sentence-final position as the reference value. Sentence context was categorized as congruent (e.g., deer-predictive when the target word was one of the two VOT continuum steps closest to "deer" or tear-predictive when the target word was one of the remaining two VOT continuum steps) or incongruent (opposite context or neutral context). The congruent sentence context was used as the reference condition in the model. The initial random effects included a random intercept of participant with random slopes of SNR, ambiguity, sentence position, and congruency. The random intercept of participant accounted for individual differences in overall absolute response times. The *buildmer* package (version 2.11; Voeten, 2023) was used to identify the best-fit model. The package evaluates each potential effect, identifying terms that contribute to the model. It then performs backward elimination using likelihood-ratio testing until arriving at a model in which all terms contribute significantly to the model. The residuals of each model were checked to ensure there were no major violations of normality or homoscedasticity.

In summary, the second analysis modeled the effects of age, sentence position, acoustic ambiguity, SNR, and their interactions on the response times of participants' phoneme categorization choices. The intercept of this model represented the response time measured for a participant at the mean age (49.2 years) in the 0 dB SNR condition when an unambiguous target word was presented at the end of a congruent context sentence. The effects of SNR, acoustic ambiguity, and sentence position were allowed to be different for each participant, and these effects were captured in the random effects.

III. RESULTS

A. Categorization performance

The data are presented as average performance for all participants at each continuum step, at each SNR, and in each sentence position (Fig. 3). Higher values on the plot represent more "deer" responses, and lower values represent more "tear" responses. Context effect can be seen as the difference between performance in the neutral sentence context (black triangles and dotted lines) compared to performance in both the "deer" and "tear" predictive sentence contexts. The effect of context on responses appears to be small until −5 dB SNR for the sentence-initial target words (Fig. 3, top row) and appears to be small but consistent across SNRs for the sentence-final target words (Fig. 3, bottom row). An analysis of the effects of SNR, continuum step, sentence position, age, and context type on the trial level data is available in the supplementary material. To best answer our research questions, we focus on the effects of the experimental factors on the context effect.

B. Context effect measures

As described in the Analysis section (Sec. II C 4), the context effect is calculated based on the difference between the total number of "deer" responses in each of the predictive context conditions compared to the total number of "deer" responses in the neutral context condition. A positive context effect would be the result of more "deer" responses in the deer-predicting context sentences than in the neutral sentences or more "tear" responses in the tear-predicting

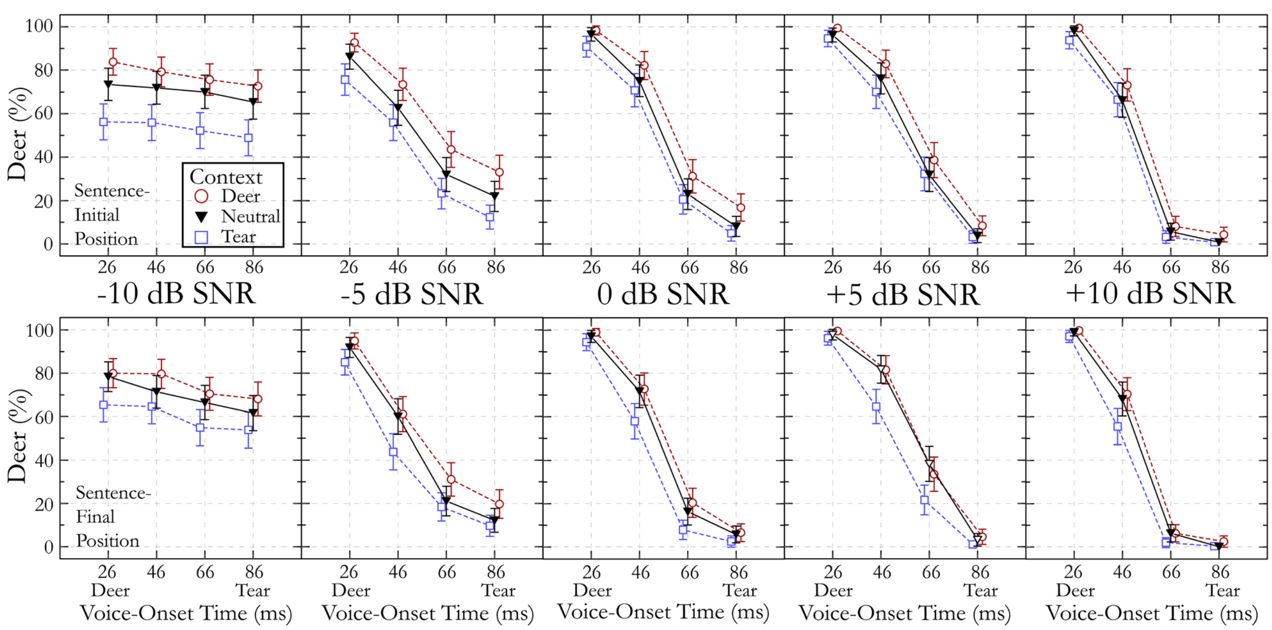

FIG. 3. Average percentage of "deer" responses (*y* axis) at each VOT along the speech continuum [*x* axis; shortest VOT ("deer" stimulus) on the left of each subplot; longest VOT ("tear" stimulus) on the right of each subplot] are shown at two target word positions within the sentence (top row: initial position; bottom row: final position), five SNRs (columns), and three sentence contexts (colors/symbols). The SNRs range from the most challenging SNR (–10 dB SNR, far left column) to the least challenging SNR (10 dB SNR, far right column). Responses are marked with red circles when the context sentence predicted "deer" as the target word, blue squares when the context sentence predicted "tear," and black triangles when the context sentence was neutral and did not predict either target word. Error bars indicate ±1 standard error of the mean.

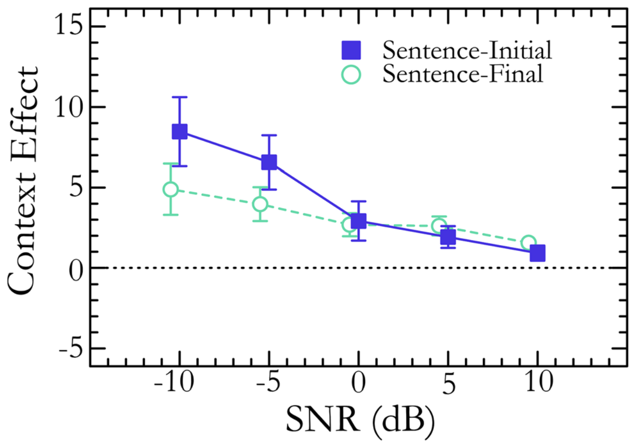

FIG. 4. The average context effect across all participants at each SNR condition for target words presented at the beginning of the context sentence (blue squares) and at the end of the context sentence (green circles). Context effect is calculated as the sum of the differences in the number of "deer" responses between the two predictive sentence contexts and the neutral sentence context. The dotted line represents performance in the neutral sentence context. Error bars indicate ±1 standard error of the mean.

context sentences than in the neutral sentences. The context effects for both deer-predictive sentences and tear-predictive sentences are combined to create an overall context effect.

Figure 4 shows the context effect at each SNR as a function of target word position collapsed across all participants. At the most difficult SNRs (i.e., −5 and −10 dB SNR), participants on average showed an increase in the effect of context. This is a logical response pattern to an increasingly inaudible target word. When the target word was presented in the sentence-initial position, participants demonstrated a greater context effect in the more difficult SNRs compared to when the target word was in the sentence-final position.

A linear MEM was created to answer our research questions about the influence of sentence position and age on the context effect for phoneme categorization of target words at various levels of degradation (SNR). The context effect for each participant was calculated at each SNR and target word position and then log-transformed to reduce the skewness of the distribution before being used as the dependent variable in the model. The final model is presented in Table II. The size of the context effect for participants was greater at SNRs < 0 dB SNR, especially for target words in the sentence-initial position (see Fig. 4) (SNR × Position-Initial, $p = 0.008$). Neither age nor any interactions that included age were statistically significant ($p > 0.05$ for all).

C. Response time measures

Response times were collected for all trials. Any response times less than 0.1 s and greater than 10 s were removed from the analysis as invalid responses. These criteria eliminated 177 trials, or 0.4% of the data. All remaining response times were transformed as described in the Analysis section (Sec. II C 4). Given that both endpoint words of the VOT continuum are clear acoustic representations of the target words while the intermediate words are acoustically ambiguous, the target words were classified as acoustically ambiguous or unambiguous for this analysis. Context was categorized as either congruent or incongruent. The incongruent context sentences included non-predictive sentences as well as sentences that predicted the opposite target word than the acoustic target word.

A linear MEM was fitted to the data with SNR, target word ambiguity, age, sentence position, and context congruency as fixed effects and SNR, sentence position, target word ambiguity, and participant as random effects (Table III). The model shows that response times were significantly predicted by various interactions between subgroupings of

TABLE II. Model output for a linear MEM fit to log-transformed derived context effect values. Reference values: Final sentence position, average SNR, and average age. Formula: Context Effect ~ SNR × Sentence Position × Age + (SNR + Sentence Position | participant).

	Context effect			
Predictors	Estimate	Standard error	*z*	*p*
(Intercept)	**2.507**	**0.043**	**57.759**	**<0.001**
SNR	−0.043	0.033	−1.282	0.206
Position-Initial	0.013	0.048	0.275	0.785
Age	−0.001	0.043	−0.021	0.983
SNR × Position-Initial	**−0.080**	**0.030**	**−2.681**	**0.008**
SNR× Age	0.015	0.034	0.430	0.669
Position-Initial × Age	0.020	0.030	0.407	0.687
SNR × Position-Initial × Age	−0.038	0.030	−1.278	0.202
Random effects	Variance	Standard deviation	Correlation	
By-participant effects	0.052	0.223		
By-participant sentence position slopes	0.052	0.228	0.36	
By-participant SNR slopes	0.025	0.158	−0.86	−0.75
Observations	360			

https://doi.org/10.1121/10.0037182

TABLE III. Model output for a linear MEM fit to log-transformed response time data. Reference values: Final sentence position, unambiguous target words, congruent sentence contexts, average SNR, and average age. Bold font denotes statistically significant factors. Formula: Response Time ∼ SNR × Sentence Position × Target Word Acoustic Ambiguity + Sentence Position × Age × (SNR + Context Congruency) + SNR × Congruency + (SNR + Sentence Position + Ambiguity | participant).

	Response time				
Predictors	Estimate	Standard error	z	p	
(Intercept)	**−0.232**	**0.059**	**−3.954**	**<0.001**	
SNR	**−0.170**	**0.018**	**−9.621**	**<0.001**	
Position-Initial	**−0.178**	**0.025**	**−7.186**	**<0.001**	
Target-Ambiguous	**0.109**	**0.009**	**12.462**	**<0.001**	
Age	**0.227**	**0.060**	**3.771**	**<0.001**	
Context-Incongruent	**0.022**	**0.006**	**3.401**	**<0.001**	
SNR × Position-Initial	**0.082**	**0.006**	**13.702**	**<0.001**	
SNR × Target-Ambiguous	**0.063**	**0.006**	**10.480**	**<0.001**	
Position-Initial × Target-Ambiguous	**−0.090**	**0.008**	**−10.592**	**<0.001**	
Position-Initial × Age	0.032	0.025	1.254	0.217	
Position-Initial × Context-Incongruent	−0.002	0.009	−0.178	0.858	
SNR × Age	−0.024	0.018	−1.353	0.185	
Age × Context-Incongruent	0.011	0.007	1.605	0.109	
SNR × Context-Incongruent	−0.008	0.004	−1.840	0.066	
SNR × Position-Initial × Target-Ambiguous	**−0.036**	**0.008**	**−4.191**	**<0.001**	
SNR × Position-Initial × Age	**−0.012**	**0.004**	**−2.791**	**0.005**	
Position-Initial × Age × Context-Incongruent	**−0.019**	**0.009**	**−2.012**	**0.044**	
Random effects	Variance	Standard deviation	Correlation		
By-participant effects	0.122	0.349			
By-participant SNR slopes	0.010	0.101	−0.41		
By-participant sentence position slopes	0.019	0.139	−0.29	−0.24	
By-participant target ambiguity slopes	0.001	0.038	−0.19	0.07	−0.14
Observations	43 023				

the fixed effects. In general, response times were shorter when the SNR was higher and when the target word occurred at the beginning of the sentence (SNR, $p < 0.001$; Position-Initial, $p < 0.001$). Response times were generally longer when the target words were acoustically ambiguous, when the context sentence was incongruent, and with increasing age (Target-Ambiguous, $p < 0.001$; Context-Incongruent, $p < 0.001$; Age, $p < 0.001$).

Response times to acoustically ambiguous target words appear to be longer than response times to acoustically unambiguous target words mainly at the more favorable SNRs. This was especially evident when the target word was at the end of the sentence [see Figs. 5(A) and 5(B)] (SNR × Position-Initial × Target-Ambiguous, $p < 0.001$). The difference between response times for target words presented in the sentence-initial vs sentence-final positions was generally larger at the more difficult SNRs (SNR × Position-Initial, $p < 0.001$). The position of the target word had a larger effect on the response times of participants in their 30s, 40s, and 50s than it did for those younger than 30 years old or older than 60 years old at most SNRs tested. At −10 dB SNR, the oldest participants showed the largest effect [see Fig. 5(C)] (SNR × Position-Initial × Age, $p < 0.01$). In Figs. 5(C) and 5(D), participants were grouped by age into decades and their performance averaged across the members of the group (5–7 participants/group) for illustration purposes. To determine if performance in the −10 dB SNR condition might be driving the three-way interactions involving SNR, the same model presented in Table III was fit to a dataset that excluded the −10 dB SNR condition. The three-way interaction between SNR, position, and target word ambiguity was no longer significant (SNR × Position-Initial × Target-Ambiguous, $p > 0.05$), but the interaction between SNR, position, and age was more significant (SNR × Position-Initial × Age, $p < 0.001$). Therefore, the −10 dB SNR condition does not appear to be the sole driver of the SNR effects in the response time analysis.

This pattern of the difference in response times based on the target word's sentence position and the participant's age is also affected by the congruency of the context sentence with the target word (Position-Initial × Age × Context-Congruent, $p < 0.05$). As in the previous interaction, participants in their 30s, 40s, and 50s showed larger differences in response times between sentence-final and sentence-initial presentations [the height of the lines in Fig. 5(D)]. The youngest and oldest participants showed the least differences due to sentence position and showed opposite effects of congruency. Younger participants showed greater position differences for congruent context sentences than incongruent context sentences [circles

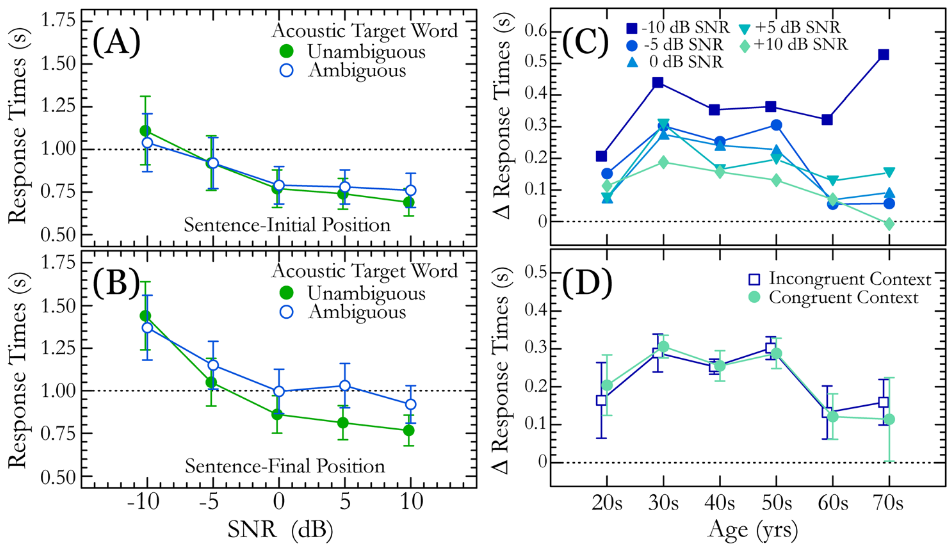

FIG. 5. Raw response times were plotted as a function of SNR when the acoustic target word was unambiguous (solid symbols) compared to ambiguous (open symbols). Target words were presented at the beginning of the sentence (A) and at the end of the sentence (B). The dotted line at 1.00 s is a reference for comparing across panels. (C) The difference in response times (sentence-final condition minus sentence-initial condition) across ages as a function of SNR. (D) The difference in response times (sentence-final condition minus sentence-initial condition) across ages as a function of context congruence. Error bars represent ±1 standard error of the mean.

higher than squares in Fig. 5(D)], while older participants showed larger position effects for incongruent context sentences than congruent sentences [squares higher than circles in Fig. 5(D)].

IV. DISCUSSION

This study tested adult participants with a wide range of ages on a phoneme categorization task where a pair of spectrally degraded target words varying in a single phoneme and embedded in background noise were presented at the beginning or end of spectrally degraded context sentences. The broad goal was to examine how sentence context interacted with age, signal degradation (specifically SNR), and sentence position to affect the perception of two contrasting speech sounds. It was hypothesized that participants would demonstrate a smaller effect of context for sentence-initial compared to sentence-final target words, that age-related differences would be seen in the response times but not in the phoneme categorization performance, and that increased age would result in a reduction of the difference between the context effects found at the beginning vs the end of a sentence.

The hypothesis for the first research question, regarding the position of the target word, was that the context effect would be greater when the target word was at the end of the sentence than when it was at the beginning of the sentence. The hypothesis was based on the context effects reported in Connine *et al.*, 1991 and Wingfield *et al.* (1994). The results of the current study, however, showed that the position of the target word interacted with context to differentially affect categorization performance in an unexpected way: there was a greater context effect when the target word was at the beginning of the sentence compared to when it was at the end of the sentence (Fig. 4).

The difference in the current findings compared to previous studies may lie in the differences in stimuli, methodology, and/or the calculation of context effects. The stimuli used in the current study were newly created and thus are inherently different from the stimuli used in either Wingfield *et al.* (1994) or Connine *et al.* (1991) (e.g., different talker, different words, different sentences). Wingfield *et al.* (1994) examined context effects by presenting a target word from the middle of a naturally spoken sentence. The target word was presented first in isolation and then with an increasing number of words from the sentence either preceding, following, or on both sides of the target word. Participants were asked to identify the target word from an open-set (i.e., no choices were provided). The number of context words was increased incrementally until the participant was able to correctly identify the target word. Context effect was calculated as the change in word recognition accuracy from the word presented in isolation. Similar to the current study, Connine *et al.* (1991) used a phonemic categorization task but calculated the context effect by comparing the total number of "dent" responses between a context sentence predicting the word "dent" and a context sentence predicting the word "tent." The method of calculating the context effect by comparing the performance on each (opposite) predictive context to performance with a neutral baseline before summing those effects to create a general context effect is unique to this study. The current method corrects for any inherent response biases on an individual basis and could be a more sensitive measure of context effects than those used previously.

The hypothesis for the second research question, regarding the effects of aging, was that age would not predict the context effect on categorization performance but would affect the response times of participants. Age-related decrements in context benefit have been shown to disappear when audibility or speech understanding performance is equated across age groups (e.g., Dubno *et al.*, 2000; Humes and Dubno, 2010; O'Neill *et al.*, 2021). This could indicate that any age-related decrements to the context benefit found in older listeners (e.g., Benichov *et al.*, 2012; Braver *et al.*, 2001; Wingfield *et al.*, 1994) were due to changes in peripheral hearing (i.e., age-related hearing loss) and not a decrement in the cognitive abilities needed to derive a context benefit. The current study found no significant effect of age on the size of the context effect (Table II), confirming the first part of the hypothesis. This finding appears to support the growing evidence that context benefit is available to listeners independent of their ages.

The second part of the hypothesis was that increasing age would correspond with longer response times to target words occurring in incongruent sentence contexts. The MEM analysis accounts for individual variability in baseline response times and possible age-related slowing—both cognitive and motor. The hypothesis was based on documented age-related deficits in the ability to inhibit irrelevant information (e.g., Federmeier and Kutas, 2005; Hasher *et al.*, 1991; Rush *et al.*, 2006; Sommers and Danielson, 1999). Table III shows that sentence position, SNR, target word ambiguity, context, and age significantly predicted response times. It is reasonable that faster response times would occur for sentence-initial target words because participants had longer to make their decision and then plan their motor response. There was a significant interaction between SNR and target word ambiguity, such that response times to the acoustically ambiguous words [Figs. 5(A) and 5(B)] were significantly slower than the response times to the acoustically unambiguous endpoint words for the easier SNRs (e.g., 10 and 5 dB SNR), but there was no difference in response times due to target word ambiguity at the more difficult SNRs. Plots of the age-related interactions with SNR, sentence position, and context show non-linear effects of age [Figs. 5(C) and 5(D)]. In general, the impact of SNR, sentence position, and context was larger for participants aged 30–60 years and smaller for the youngest and oldest participants. The reason for this pattern is unclear. It may reflect response strategies that differed in the group of participants aged 30–60 years from the youngest and oldest participants. Although age significantly affected response times on the current task, the exact relationship between age and response times as a measure of context effect remains uncertain.

The hypothesis for the third research question, regarding the interaction between sentence position and age on the derived context effects or response times, was that participants would show smaller differences in the context effects or response times between the two sentence positions with increasing age. Smaller context effects have been found on an open-set word recognition task for older listeners with sentence-initial target words compared to younger listeners, but not for sentence-final target words (Wingfield *et al.*, 1994). The current study found no significant interaction between age and sentence position on the size of the context effect for categorization of phonemes in contrasting word pairs (Table II). However, there were significant interactions between age and sentence position in the response time analysis, although these appeared only in three-way interactions involving either SNR or context. Participants in the middle of the age range showed the largest effects of sentence position. The oldest and youngest participants showed smaller effects of sentence position that depended on the SNR condition and the context sentence [Figs. 5(C) and 5(D)]. Thus, results from the current study only partially supported the hypothesis. Older listeners showed smaller differences in the context effects on response time measures during phoneme categorization of contrasting target words between the two sentence positions when compared with middle-aged listeners, but not when compared with listeners in their 20s.

The predicted age-related inhibition deficit, which would have been shown by a significant interaction between age and congruency, was not found in the phoneme categorization performance, but was present in the response times (significant interaction between sentence position, age, and congruency). This lack of a significant interaction between age and congruency in the derived context effect could indicate that there is no significant difference in the use of context with increasing age when the speech is spectrally degraded and the target words are partially obscured by background noise. The association between age and response times remains unclear. Alternatively, the lack of a significant interaction between age and congruency in the derived context effect could indicate that older adults use compensatory strategies, such as leveraging their knowledge of vocabulary and word predictability, to overcome any age-related speech processing deficits and to match the performance of younger adults (as in Dubno *et al.*, 2000; O'Neill *et al.*, 2021; Pichora-Fuller, 2008). More research is needed to fully understand the effect of age on context usage.

This study may have implications for people who are constantly listening to spectrally degraded speech, such as those who use cochlear implants. Spectrally degraded representations of sound are conveyed by cochlear-implant sound processors to the auditory nerves of listeners. Understanding how aging affects the ability to use context cues is vitally important for listeners with cochlear implants, given that many of them are currently over 65 years (e.g., Başkent *et al.*, 2016; Schvartz-Leyzac *et al.*, 2025; Shader *et al.*, 2020; Yang and Cosetti, 2016). The finding that unclear target words in sentence-initial positions cause listeners with normal hearing to interpret the word as congruent with the context more than when the target words are in sentence-final positions, could be indicative of the effortful nature of understanding spectrally degraded speech. If faster response times reflect ease of listening in the current study, the youngest and oldest listeners demonstrated less effort than the listeners aged 30–60 years. Young listeners might show low listening effort because of their rapid speed of processing, and older listeners might show low listening effort because of their compensatory strategies, such as a reliance on context cues. As speech unfolds over time, a common strategy is to rely on the surrounding context, rather than on the degraded acoustic cues, to determine an unclear word's identity. Given that a conversation links sentences together, listeners are unlikely to have the time to focus on acoustic

https://doi.org/10.1121/10.0037182

cues to the extent they were able to in the sentence-final presentations in the current study. Thus, this study likely underestimates the context effect that is experienced by listeners with normal hearing or with cochlear implants in typical conversations.

Future studies to continue to explore the interaction between age and perception of degraded speech should either test participants with normal hearing who have adapted to degraded speech, rather than presenting degraded speech in an acute manner, or test participants who use cochlear implants every day. Perhaps other methodologies would provide more insight into the individual strategies employed by listeners of different ages. For example, a more objective measure of real-time language processing, such as gaze tracking in a visual world paradigm, may provide researchers with insight into the listener's decision-making process as it unfolds over time, and evaluate the impact of sentence position, SNR, and VOT at a finer-grained level (e.g., Ben-David *et al.*, 2011; Cooper, 1974; Harel-Arbeli *et al.*, 2021). Future studies could also explore additional phonemic contrasts in various word pairs to assess the generalizability of these findings.

V. CONCLUSION

Acoustic-hearing listeners who categorized contrasting initial phonemes varying in VOT in target words that were presented in various sentence positions, semantic contexts, and levels of background noise showed some expected, and some unexpected, patterns of context benefit. As expected, the level of background noise, the position of the target word, and the congruency of the sentence with the target word all significantly affected performance. Contrary to expectations, larger context effects were seen for target words presented at the beginning of sentences than for target words presented at the end of sentences at the most difficult SNRs of −10 and −5 dB. Also, contrary to expectations, the effects of background noise level, target word position, and context sentence-target word congruency on response times obtained during the phoneme categorization task were larger for participants in the middle of the age range and smaller for the youngest and oldest participants tested. Therefore, older listeners appear to be able to use lexical context to inform their interpretation of unclear words in a similar manner to younger listeners when speech is degraded, possibly using strategies that are only employed by acoustic hearing participants over 60 years old. This study supports the use of context cues to aid understanding of degraded speech for adults of all ages.

SUPPLEMENTARY MATERIAL

See the supplementary material for an analysis of the raw phoneme categorization data.

ACKNOWLEDGMENTS

This work was supported by NIH-NIDCD Grant Nos. R01 DC020316 (Goupell) and T32 DC000046 (Comparative and Evolutionary Biology of Hearing training grant; consecutive PIs: Gordon-Salant, Goupell, trainee: Tinnemore). Portions of this work were presented at the 183rd Meeting of the Acoustical Society of America. We would like to thank Pallavi Atluri, Seren Avenilla, Erin Doyle, Milana Klopouh, and Vanessa Reyes for help in data collection; and Stefanie Kuchinsky for insight and assistance with the statistical analysis.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Ethics Approval

The experimental procedures described in this article were approved by the University of Maryland Institutional Review Board for Human Subjects Research. Participants were recruited via word of mouth through lab members and via a database of participants who had previously consented to be contacted with future research opportunities. Participants provided written informed consent and were paid for their time.

DATA AVAILABILITY

The data that support the findings of this study are openly available in Open Science Framework at https://osf.io/ky6wa.

Amichetti, N. M., Atagi, E., Kong, Y.-Y., and Wingfield, A. (**2018**). "Linguistic context versus semantic competition in word recognition by younger and older adults with cochlear implants," Ear Hear. **39**, 101–109.

ANSI (**2018**). ANSI S3.6-2018, *American National Standard Specification for Audiometers* (American National Standards Institute, New York).

Bacon, S. P., and Healy, E. W. (**2000**). "Effects of ipsilateral and contralateral precursors on the temporal effect in simultaneous masking with pure tones," J. Acoust. Soc. Am. **107**, 1589–1597.

Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (**2013**). "Random effects structure for confirmatory hypothesis testing: Keep it maximal," J. Mem. Lang. **68**, 255–278.

Başkent, D., Clarke, J., Pals, C., Benard, M. R., Bhargava, P., Saija, J., Sarampalis, A., Wagner, A., and Gaudrain, E. (**2016**). "Cognitive compensation of speech perception with hearing impairment, cochlear implants, and aging: How and to what degree can it be achieved?," Trends Hear. **20**, 2331216516670279.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (**2015**). "Fitting linear mixed-effects models using lme4," J. Stat. Softw. **67**(1), 1–48.

Ben-David, B. M., Chambers, C. G., Daneman, M., Pichora-Fuller, M. K., Reingold, E. M., and Schneider, B. A. (**2011**). "Effects of aging and noise on real-time spoken word recognition: Evidence from eye movements," J. Speech Lang. Hear. Res. **54**, 243–262.

Ben-David, B. M., Erel, H., Goy, H., and Schneider, B. A. (**2015**). " 'Older is always better': Age-related differences in vocabulary scores across 16 years," Psychol. Aging **30**, 856–862.

Benichov, J., Cox, L. C., Tun, P. A., and Wingfield, A. (**2012**). "Word recognition within a linguistic context: Effects of age, hearing acuity, verbal ability and cognitive function," Ear Hear. **33**, 250–256.

Bhandari, P., Demberg, V., and Kray, J. (**2021**). "Semantic predictability facilitates comprehension of degraded speech in a graded manner," Front. Psychol. **12**, 714485.

Bilger, R. C., Nuetzel, J. M., Rabinowitz, W. M., and Rzeczkowski, C. (**1984**). "Standardization of a test of speech perception in noise," J. Speech Lang. Hear. Res. **27**, 32–48.

Boothroyd, A., and Nittrouer, S. (**1988**). "Mathematical treatment of context effects in phoneme and word recognition," J. Acoust. Soc. Am. **84**, 101–114.

Braver, T. S., Barch, D. M., Keys, B. A., Carter, C. S., Cohen, J. D., Kaye, J. A., Janowsky, J. S., Taylor, S. F., Yesavage, J. A., Mumenthaler, M. S., Jagust, W. J., and Reed, B. R. (**2001**). "Context processing in older adults: Evidence for a theory relating cognitive control to neurobiology in healthy aging," J. Exp. Psychol. Gen. **130**, 746–763.

Connine, C. M. (**1987**). "Constraints on interactive processes in auditory word recognition: The role of sentence context," J. Mem. Lang. **26**, 527–538.

Connine, C. M., Blasko, D. G., and Hall, M. (**1991**). "Effects of subsequent sentence context in auditory word recognition: Temporal and linguistic constraints," J. Mem. Lang. **30**, 234–250.

Connine, C. M., and Clifton, C. (**1987**). "Interactive use of lexical information in speech perception," J. Exp. Psychol. Hum. Percept. Perform. **13**, 291–299.

Cooper, R. M. (**1974**). "The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing," Cogn. Psychol. **6**, 84–107.

Cruickshanks, K. J., Wiley, T. L., Tweed, T. S., Klein, B. E. K., Klein, R., Mares-Perlman, J. A., and Nondahl, D. M. (**1998**). "Prevalence of hearing loss in older adults in Beaver Dam, Wisconsin: The epidemiology of hearing loss study," Am. J. Epidemiol. **148**, 879–886.

Cychosz, M., Winn, M. B., and Goupell, M. J. (**2024**). "How to vocode: Using channel vocoders for cochlear-implant research," J. Acoust. Soc. Am. **155**, 2407–2437.

Ding, N., Melloni, L., Zhang, H., Tian, X., and Poeppel, D. (**2016**). "Cortical tracking of hierarchical linguistic structures in connected speech," Nat. Neurosci. **19**, 158–164.

Dorman, M. F., Loizou, P. C., and Rainey, D. (**1997**). "Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs," J. Acoust. Soc. Am. **102**, 2403–2411.

Dubno, J. R., Ahlstrom, J. B., and Horwitz, A. R. (**2000**). "Use of context by young and aged adults with normal hearing," J. Acoust. Soc. Am. **107**, 538–546.

Dubno, J. R., Dirks, D. D., and Morgan, D. E. (**1984**). "Effects of age and mild hearing loss on speech recognition in noise," J. Acoust. Soc. Am. **76**, 87–96.

Federmeier, K. D. (**2007**). "Thinking ahead: The role and roots of prediction in language comprehension," Psychophysiology **44**, 491–505.

Federmeier, K. D., and Kutas, M. (**2005**). "Aging in context: Age-related changes in context use during language comprehension," Psychophysiology **42**, 133–141.

Fletcher, A., McAuliffe, M., Kerr, S., and Sinex, D. (**2019**). "Effects of vocabulary and implicit linguistic knowledge on speech recognition in adverse listening conditions," Am. J. Audiol. **28**, 742–755.

Friesen, L. M., Shannon, R. V., Başkent, D., and Wang, X. (**2001**). "Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants," J. Acoust. Soc. Am. **110**, 1150–1163.

Golestani, N., Rosen, S., and Scott, S. K. (**2009**). "Native-language benefit for understanding speech-in-noise: The contribution of semantics," Bilingualism **12**, 385–392.

Gordon-Salant, S., and Cole, S. S. (**2016**). "Effects of age and working memory capacity on speech recognition performance in noise among listeners with normal hearing," Ear Hear. **37**, 593–602.

Gordon-Salant, S., Frisina, R. D., Fay, R. R., and Popper, A. N. (**2010**). *The Aging Auditory System* (Springer Science & Business Media, New York), Vol. 34.

Hannemann, R., Obleser, J., and Eulitz, C. (**2007**). "Top-down knowledge supports the retrieval of lexical information from degraded speech," Brain Res. **1153**, 134–143.

Harel-Arbeli, T., Wingfield, A., Palgi, Y., and Ben-David, B. M. (**2021**). "Age-related differences in the online processing of spoken semantic context and the effect of semantic competition: Evidence from eye gaze," J. Speech Lang. Hear. Res. **64**, 315–327.

Hasher, L., Stoltzfus, E. R., Zacks, R. T., and Rypma, B. (**1991**). "Age and inhibition," J. Exp. Psychol. Learn. Mem. Cogn. **17**, 163–169.

Helfer, K. S., and Wilber, L. A. (**1990**). "Hearing loss, aging, and speech perception in reverberation and noise," J. Speech Lang. Hear. Res. **33**, 149–155.

Humes, L. E., Busey, T. A., Craig, J., and Kewley-Port, D. (**2013**). "Are age-related changes in cognitive function driven by age-related changes in sensory processing?," Atten. Percept. Psychophys. **75**, 508–524.

Humes, L. E., and Dubno, J. R. (**2010**). "Factors affecting speech understanding in older adults," in *Aging Auditory Systems: Springer Handbook of Auditory Research*, edited by S. Gordon-Salant, R. D. Frisina, A. N. Popper, and R. R. Fay (Springer, New York), pp. 211–257.

Jurafsky, D. (**1996**). "A probabilistic model of lexical and syntactic access and disambiguation," Cogn. Sci. **20**, 137–194.

Kalikow, D. N., Stevens, K. N., and Elliott, L. L. (**1977**). "Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability," J. Acoust. Soc. Am. **61**, 1337–1351.

Kavé, G., and Halamish, V. (**2015**). "Doubly blessed: Older adults know more vocabulary and know better what they know," Psychol. Aging **30**, 68–73.

Liberman, A. M., Cooper, F. S., Shankweiler, D., and Studdert-Kennedy, M. (**1967**). "Perception of the speech code," Psychol. Rev. **74**, 431–461.

Liberman, A. M., Repp, B. H., Eccardt, T., and Pesetsky, D. (**1977**). "Some relations between duration of silence and duration of friction noise as joint cues for fricatives, affricates, and stops," J. Acoust. Soc. Am. **62**, S78.

Lin, F. R. (**2011**). "Hearing loss and cognition among older adults in the United States," J. Gerontol. A. Biol. Sci. Med. Sci. **66A**, 1131–1136.

Lipnicki, D. M., Crawford, J. D., Dutta, R., Thalamuthu, A., Kochan, N. A., Andrews, G., Lima-Costa, M. F, Castro-Costa, E., Brayne, C., Matthews, F. E., Stephan, C. M., Lipton, R. B., Katz, M. J., Ritchie, K., Scali, J., Ancelin, M.-L., Scarmeas, N., Yannakoulia, M., Dardiotis, E., Lam, L. C. W., Wong, C. H. Y., Fung, A. W. T., Guaita, A., Vaccaro, R., Davin, A., Kim, K. W., Han, J. W., Kim, T. H., Anstey, K. J., Cherbuin, N., Butterworth, P., Scazufca, M., Kumagai, S., Chen, S., Narazaki, K., Ng, T. P., Gao, Q., Reppermund, S., Brodaty, H., Lobo, A., Lopez-Anton, R., Santabárbara, J., and Sachdev, P. S. (**2017**). "Age-related cognitive decline and associations with sex, education and apolipoprotein E genotype across ethnocultural groups and geographic regions: A collaborative cohort study," PLoS Med. **14**, e1002261.

Marslen-Wilson, W., and Tyler, L. K. (**1980**). "The temporal structure of spoken language understanding," Cognition **8**, 1–71.

McClelland, J. L., and Elman, J. L. (**1986**). "The TRACE model of speech perception," Cognit. Psychol. **18**, 1–86.

Milburn, E., Dickey, M. W., Warren, T., and Hayes, R. (**2023**). "Increased reliance on world knowledge during language comprehension in healthy aging: Evidence from verb-argument prediction," Aging Neuropsychol. Cogn. **30**, 1–33.

Miller, G. A., Heise, G. A., and Lichten, W. (**1951**). "The intelligibility of speech as a function of the context of the test materials," J. Exp. Psychol. **41**, 329–335.

Morton, J. (**1969**). "Interaction of information in word recognition," Psychol. Rev. **76**, 165–178.

Obleser, J., and Kotz, S. A. (**2011**). "Multiple brain signatures of integration in the comprehension of degraded speech," Neuroimage **55**, 713–723.

Obleser, J., Wise, R. J. S., Dresner, M. A., and Scott, S. K. (**2007**). "Functional integration across brain regions improves speech perception under adverse listening conditions," J. Neurosci. **27**, 2283–2289.

O'Neill, E. R., Parke, M. N., Kreft, H. A., and Oxenham, A. J. (**2021**). "Role of semantic context and talker variability in speech perception of cochlear-implant users and normal-hearing listeners," J. Acoust. Soc. Am. **149**, 1224–1239.

Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., and Smith, P. K. (**2002**). "Models of visuospatial and verbal memory across the adult life span," Psychol. Aging **17**, 299–320.

Pichora-Fuller, M. K. (**2008**). "Use of supportive context by younger and older adult listeners: Balancing bottom-up and top-down information processing," Int. J. Audiol. **47**, S72–S82.

Pichora-Fuller, M. K., Schneider, B. A., and Daneman, M. (**1995**). "How young and old adults listen to and remember speech in noise," J. Acoust. Soc. Am. **97**, 593–608.

R Core Team (**2024**). "R: A language and environment for statistical computing," https://www.R-project.org/.

Repp, B. H., Liberman, A. M., and Harnad, S. (**1984**). "Phonetic category boundaries are flexible," Haskins Laboratories Status Report on Speech Research (SR-77/78), pp. 31–53.

Rush, B. K., Barch, D. M., and Braver, T. S. (**2006**). "Accounting for cognitive aging: Context processing, inhibition or processing speed" Aging Neuropsychol. Cogn. **13**, 588–610.

Saija, J. D., Akyürek, E. G., Andringa, T. C., and Başkent, D. (**2014**). "Perceptual restoration of degraded speech is preserved with advancing age," J. Assoc. Res. Otolaryngol. **15**, 139–148.

Schvartz-Leyzac, K. C., McClaskey, C. M., Dias, J. W., Pfingst, B. E., and Harris, K. C. (**2025**). "Contributions of auditory nerve density and synchrony to speech understanding in older cochlear implant users," J. Assoc. Res. Otolaryngol. **26**, 317–329.

Shader, M. J., Gordon-Salant, S., and Goupell, M. J. (**2020**). "Impact of aging and the electrode-to-neural interface on temporal processing ability in cochlear-implant users: Amplitude-modulation detection thresholds," Trends Hear. **24**, 2331216520936160.

Sohoglu, E., Peelle, J. E., Carlyon, R. P., and Davis, M. H. (**2012**). "Predictive top-down integration of prior knowledge during speech perception," J. Neurosci. **32**, 8443–8453.

Sommers, M. S., and 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.

Stroop, J. R. (**1935**). "Studies of interference in serial verbal reactions," J. Exp. Psychol. **18**, 643–662.

Tinnemore, A. R., Gordon-Salant, S., and Goupell, M. J. (**2020**). "Audiovisual speech recognition with a cochlear implant and increased perceptual and cognitive demands," Trends Hear. **24**, 2331216520960601.

Tulsky, D. S., Carlozzi, N., Chiaravalloti, N. D., Beaumont, J. L., Kisala, P. A., Mungas, D., Conway, K., and Gershon, R. (**2014**). "NIH Toolbox Cognition Battery (NIHTB-CB): List sorting test to measure working memory," J. Int. Neuropsychol. Soc. **20**, 599–610.

Verhaeghen, P. (**2003**). "Aging and vocabulary score: A meta-analysis," Psychol. Aging **18**, 332–339.

Voeten, C. C. (**2023**). "buildmer: Stepwise elimination and term reordering for mixed-effects regression," https://CRAN.R-project.org/package=buildmer.

Wild, C. J., Yusuf, A., Wilson, D. E., Peelle, J. E., Davis, M. H., and Johnsrude, I. S. (**2012**). "Effortful listening: The processing of degraded speech depends critically on attention," J. Neurosci. **32**, 14010–14021.

Wingfield, A., Alexander, A. H., and Cavigelli, S. (**1994**). "Does memory constrain utilization of top-down information in spoken word recognition? Evidence from normal aging," Lang. Speech **37**, 221–235.

Winn, M. B., and Litovsky, R. Y. (**2015**). "Using speech sounds to test functional spectral resolution in listeners with cochlear implants," J. Acoust. Soc. Am. **137**, 1430–1442.

Wotton, J. M., Elvebak, R. L., Moua, L. C., Heggem, N. M., Nelson, C. A., and Kirk, K. M. (**2011**). "Congruent and incongruent semantic context influence vowel recognition," Lang. Speech **54**, 341–360.

Yang, Z., and Cosetti, M. (**2016**). "Safety and outcomes of cochlear implantation in the elderly: A review of recent literature," J. Otol. **11**, 1–6.